AD-A062 651 ANALYTICAL SYSTEMS ENGINEERING CORP BURLINGTON MA F/G 9/2

MATHEMATICAL ANALYSIS AND SOFTWARE IMPLEMENTATION FOR PHYSICAL --ETC(U)

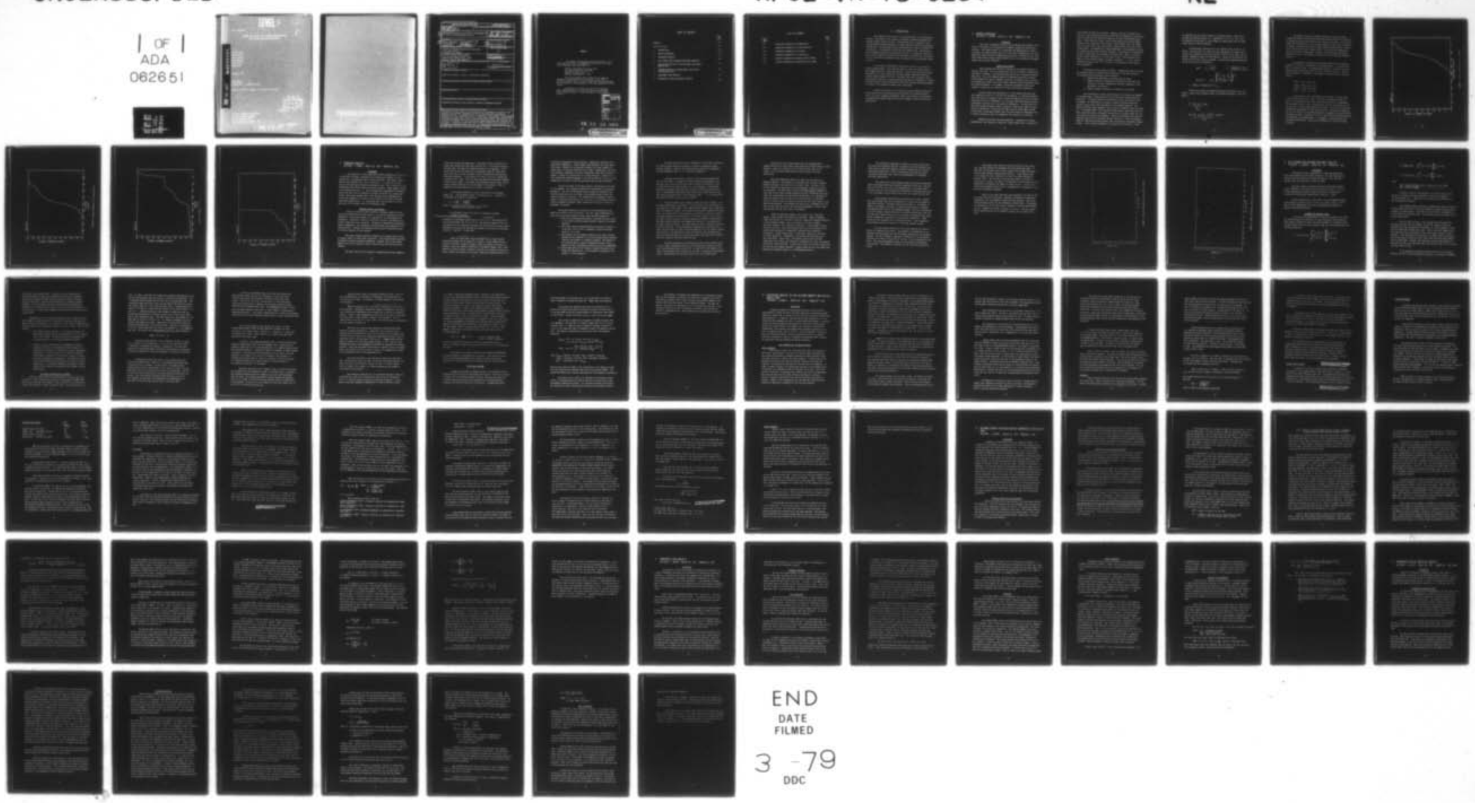
NOV 78 E J ZIEMBA, A LACROIX, J CROSBY F19628-78-C-0157

UNCLASSIFIED AFGL-TR-78-0254 NL

1 OF 1

ADA

062651



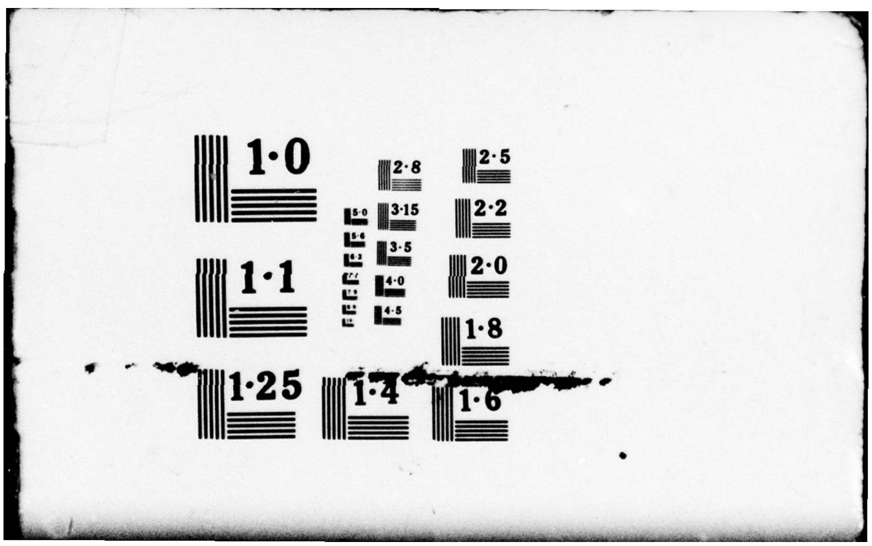
END

DATE

FILMED

3-79

DDC


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5
5·0

LEVEL II

(13)

AFGL-TR-78-0254

MATHEMATICAL ANALYSIS AND SOFTWARE IMPLEMENTATION FOR PHYSICAL AND ENGINEERING DATA

Edward Ziemba
Kathleen Lewis
Andre LaCroix
Jean Tremblay
John Crosby
Peter Chin

November 1978

Final Report
16 May 1978 - 30 September 1978

Approved for Public Release; Distribution Unlimited

Air Force Geophysics Laboratory
Air Force Systems Command
United States Air Force
Hanscom AFB, Massachusetts 01731

78 12 13 006

Qualified requesters may obtain additional copies from the Defense Documentation Center. All others should apply to the National Technical Information Service.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 18 19 AFGL-TR-78-0254	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) 6 Mathematical Analysis and Software Implementation for Physical and Engineering Data	5. TYPE OF REPORT & PERIOD COVERED 9 FINAL rept. 16 May 78 - 30 Sept 78
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Edward J. Ziemba, Kathleen Lewis, Andre LaCroix, Peter Chin, John Crosby, Jean Tremblay	8. CONTRACT OR GRANT NUMBER(s) 15 F19628-78-C-0157
9. PERFORMING ORGANIZATION NAME AND ADDRESS Analytical Systems Engineering Corporation, 5 Old Concord Road, Burlington, MA 01803	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS N/A
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts 01731, Monitor: John F. Kellaher/SUA	12. REPORT DATE 11 November 1978
	13. NUMBER OF PAGES 75
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 75 p.	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: Distribution Unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Mathematical Analysis, Data Analysis, Computer Programming Problems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this report is to describe most of mathematical analysis, data analysis, and computer programming problems performed under Contract No. F19628-78-C-0157. The problems vary in complexity from straight forward program adaptation to tasks requiring analysis, determining, implementing, and sometimes deriving algorithms best suited to perform the calculations. The problems are discussed in summary form and analysis and programming techniques required are outlined. It should be noted that most of these problems are a continuation of the efforts undertaken during the period of Contract No. F19628-77-C-0174.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE 410 174

2

PREFACE

The analysis and computer programs described in this report are the results of mathematical, analytical, and software development tasks performed for:

Analysis and Simulation Branch (SUA)
Air Force Geophysics Laboratory
Hanscom Air Force Base
Bedford, Massachusetts 01731

The writers express their thanks to Mr. Robert E. McInerney, Branch Chief, and in particular to Mr. John F. Kellaher, Contract Monitor, whose support and technical guidance were invaluable in the performing of the tasks described in this report.

In addition, the authors would like to thank the AFGL research personnel for whom the tasks were performed. Their cooperation and technical assistance is truly appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

The purpose of this report is to describe most of the mathematical analysis, data analysis and computer programming problems performed under Contract Number F19628-78-C-0157. The problems vary in complexity from straightforward program adaptation to tasks requiring analysis, determining, implementing, and sometimes deriving algorithms best suited to perform the calculations. The problems are discussed in summary form. The analysis and programming techniques required are outlined. It should be noted that all of these problems, except the one described in section two, are continuation of the efforts undertaken during the period of Contract Number F19628-77-C-0174.

In the subject description, all but one task are referenced by a single problem number. This task was addressed under two requests but is best summarized with a single description. Further, most of the tasks were described in the final reports for Contract Number F19628-76-C-0203 and Contract Number F19628-77-C-0174. To present the total functional flow of the problems, portions of items previously described are reiterated. Differentiation between recent modifications and total software packages developed is stated in the individual summary where applicable.

Software described in this report have been documented separately. These programs can be obtained from The Analysis and Simulation Branch (SUA), Air Force Geophysics Laboratory (AFGL) upon request by referencing the appropriate problem numbers and project numbers listed with each task description.

2. MEPPEN CLIMATOLOGY

INITIATOR: E. BERTONI PROJECT NO: 6670 PROBLEM NO: 4003

BACKGROUND

This task is designed to assist the Climatology and Dynamics Branch (LYD) in the study of the climatology of Meppen, Germany and surrounding stations. The desired result is to determine how representative a year's collection of surface weather data is of long term climatology for the area. Three data tapes containing climatological data are available for the study. These tapes contain hourly surface observation at five stations near Meppen, Germany for the years 1968-1976.

ANALYSIS AND RESULTS

Three programs were written to assist in the analysis required to complete this task. The first program was written to extract, reformat and save data of interest. The second program was written to calculate monthly statistics. The third program combines the monthly statistics to obtain either seasonal per year or seasonal statistics accumulated over all years. All programs are written such that further analysis could be performed using the existing software. For example, additional parameters could be extracted or statistics as a function of time of day could be generated.

The data tapes were generated on a 32 bit machine and each tape was composed of either 16 or 32 bits of information. A conversion program (CVP) was written to unpack 60 bits of input information into a 60 bit word containing 16 bits of information. In this way, each field could be easily represented by either one or two words of information. All data was in integer form unless otherwise stated, therefore, there was no need for floating point conversion. Since four CDC 6600 sixty-bit words equal fifteen 16-bit words, the program converts four CDC input words of information into 15 words containing 16 bits of information with zero-fill on the left.

Problems arose with the data extraction. According to the data documentation, the variable in field C01, record length, was to contain the total number of bytes in the record. However, this was not found to be the case. Since the block size and record size are variables, it was necessary to search for the records by keying on an easily identifiable parameter. Since the first record of each block always has the same length, except for the "additional Data Section," which appears after the needed data, the year, month, day, hour could be obtained from the first record of each block. By searching for these parameters in the subsequent record, all records in each block could be obtained. The variable M in program CVP is used to keep track of the number of records read. When it is time to buffer in more information, M is set back to zero. The additional parameters, such as station number, latitude, longitude, . . . etc., can be obtained by deleting the "C" in Column 1 of the statements that extract this information.

This data was then stored on separate, reformatted tapes and listed. The analysis which was required after completion of the data extraction in the proper units consisted of the following:

1) determine how much data is missing from the tapes,
2) compute cumulative frequency distribution for each month for visibility, ceiling, temperature, dew point depression, wind direction, and velocity,
3) compute seasonal variations of visibility and ceiling.

After preliminary outputs were obtained, the software was expanded to calculate means, standard deviations, and frequency distributions for each month. A variable array was incorporated in the calculation of the histograms. An algorithm was included to process multiple hourly data records, and all hours were considered in the expanded analysis. The monthly statistics were generated, and written on a magnetic tape for future use. The program BERT 1 which generates these monthly statistics first defines, partially within the program and partially via input cards, the various ranges of interest per parameter. Second, the program initializes all temporary arrays used to determine monthly statistics. Then, the program reads the tape which contains weather data from stations surrounding Meppen, Germany. (Each record corresponds to a particular hour and day.) Data is compared with the various ranges of legitimate values. The data is divided into various ranges, e.g., temperature < 0°F, 0-4°, 4-9°, etc. The number of occurrences in each category is saved along with the legitimate values in various matrices.

This procedure continues until one complete month of data is processed. Once the data for a month has been processed, the monthly statistics are calculated, values are printed and written onto tape, and the monthly, temporary arrays are reinitialized. Proces ing continues for all months until an end of file is reached. The means and standard deviations per matrix location are defined by

$$\text{Mean} = \quad \overline{P}_{ij} = \frac{\sum_{j=1}^{n} P_{ij}}{n}$$

P_{ij} - observation for parameter i, range j

$$\text{Std. Dev.} = \quad Sd_{ij} = \left(\frac{n \sum_{i=1}^{n} P_{ij}^2 - \overline{P}_{ij}^2}{n\,(n-1)} \right)^{1/2}$$

n - number of observations of P_{ij}

These matrix means and standard deviations are saved and are combined to obtain total parameter means and standard deviations. For parameter i:

$$\overline{P}_i = \frac{\Sigma \overline{P}_{ij} \, X n_{ij}}{\Sigma n_{ij}} = \text{mean}$$

$$Sd_i = \left(\frac{\sum_j \left(n_{ij} \, \overline{P}_{ij} - \overline{P}_i^{\,2} \right)}{\Sigma n_{ij}} \right)^{1/2} = \text{std dev.}$$

The monthly count per parameter array can be used to produce probability density functions and cumulative density functions. Statistics as described above are calculated on all ranges of interest. These values along with the monthly means and standard deviations are saved. The monthly summary tapes were then used to obtain seasonal statistics on a yearly basis, and for the entire 9 years of data. All hours were combined to obtain the statistics generated. However, future analysis may require analysis per hour. The software was written to facilitate this expanded analysis.

To calculate the seasonal distributions desired for any station, program CYM is used in conjunction with program BERT 1. The output from BERT 1 is used as input. The program first defines, partially via input and partially within source code, the cell widths of interest. Then the data from BERT 1 is processed. The seasonal definition used is the following:

Winter - Dec., Jan., Feb.

Spring - March, April, May

Summer - June, July, Aug.

Fall - Sept., Oct., Nov.

Distributions for six parameters are obtained. The data read is available one month at a time. This data is labelled by year and month. If distributions per year are desired, then data is accumulated until December data is read. Distributions are then calculated, printed and plotted. If total distributions are desired, they are calculated upon reaching the end-of-file on the input data tape. Six plots are obtained for each run for the six variable parameters. Examples of the plots being produced are shown in Figures 2-1 to 2-4. Some of the statistics and plots are generated using coded values. For example, ceiling statistics are calculated using levels. Each level corresponds to a range of heights in meters.

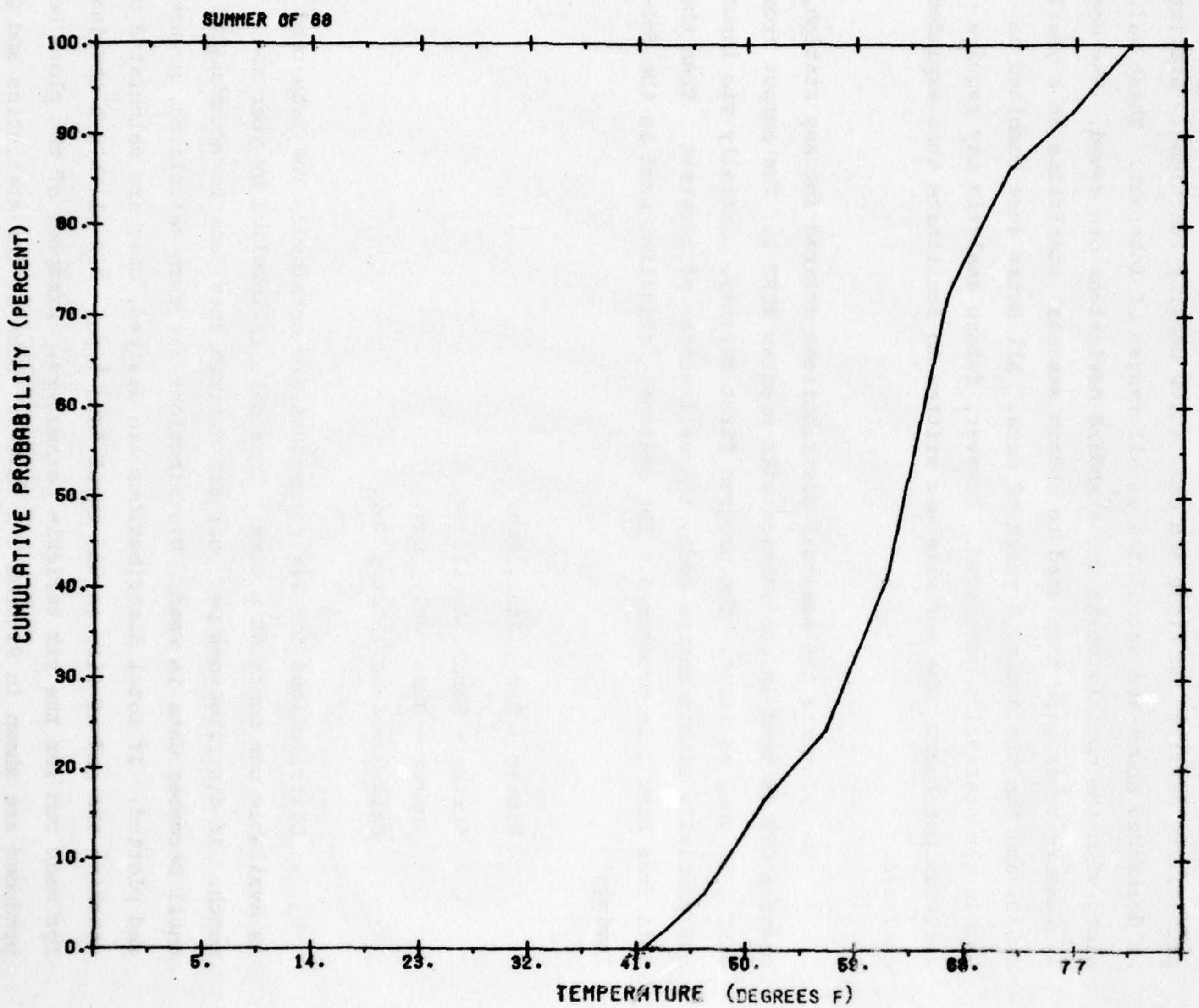


FIGURE 2-1 CUMULATIVE PROBABILITY OF TEMPERATURE

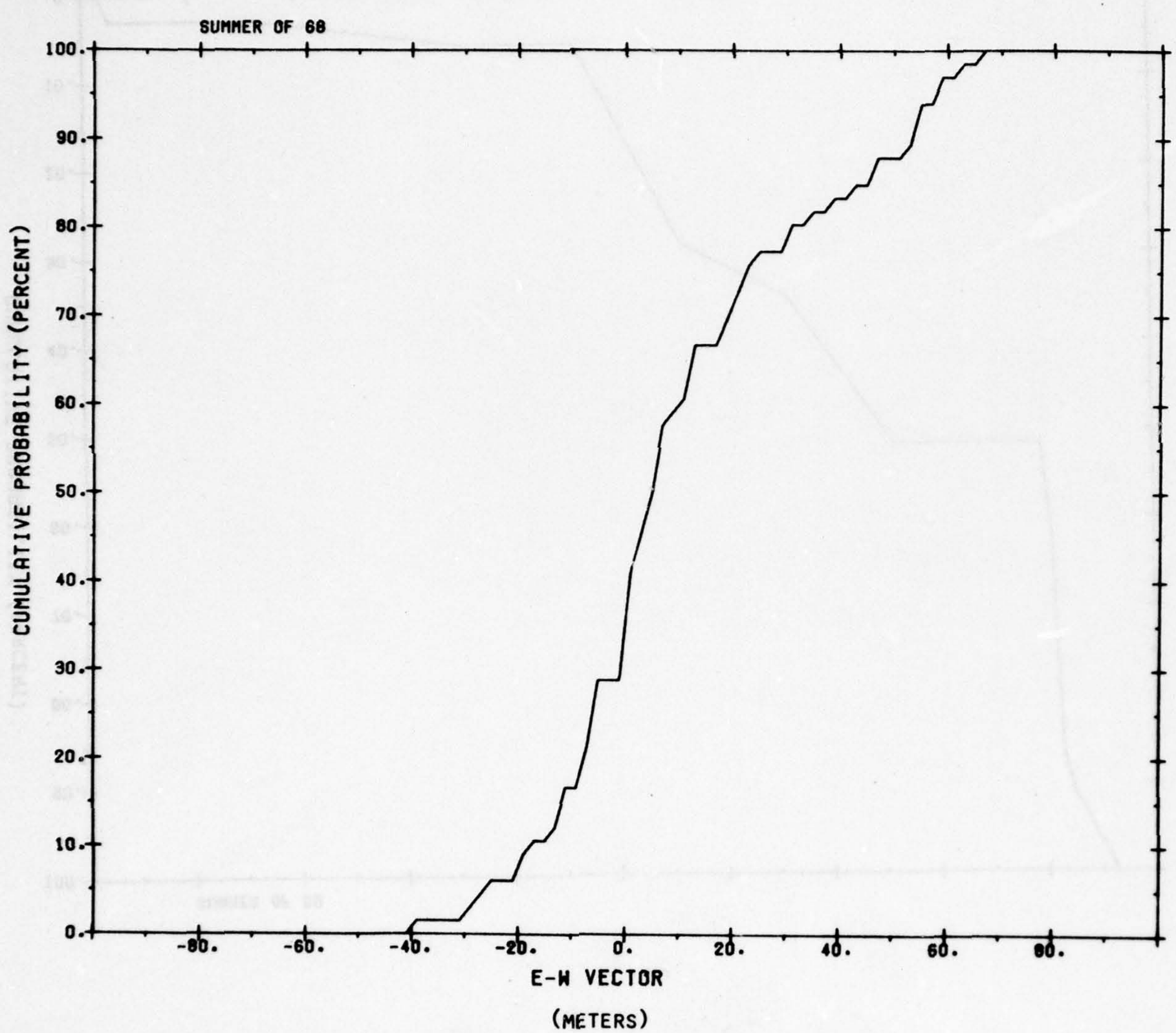


FIGURE 2-2 CUMULATIVE PROBABILITY OF WIND DIRECTION

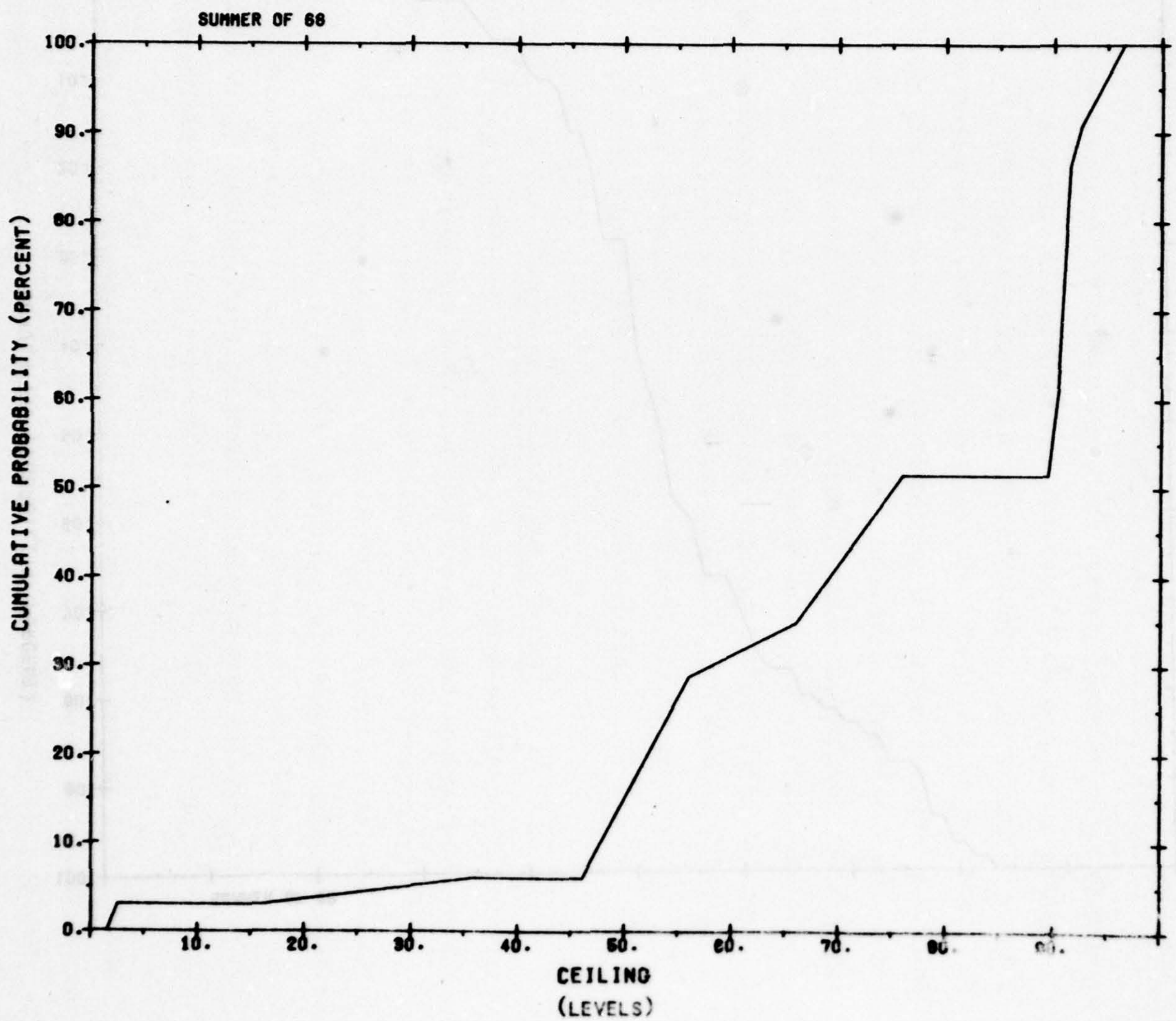


FIGURE 2-3 CUMULATIVE PROBABILITY OF CEILING

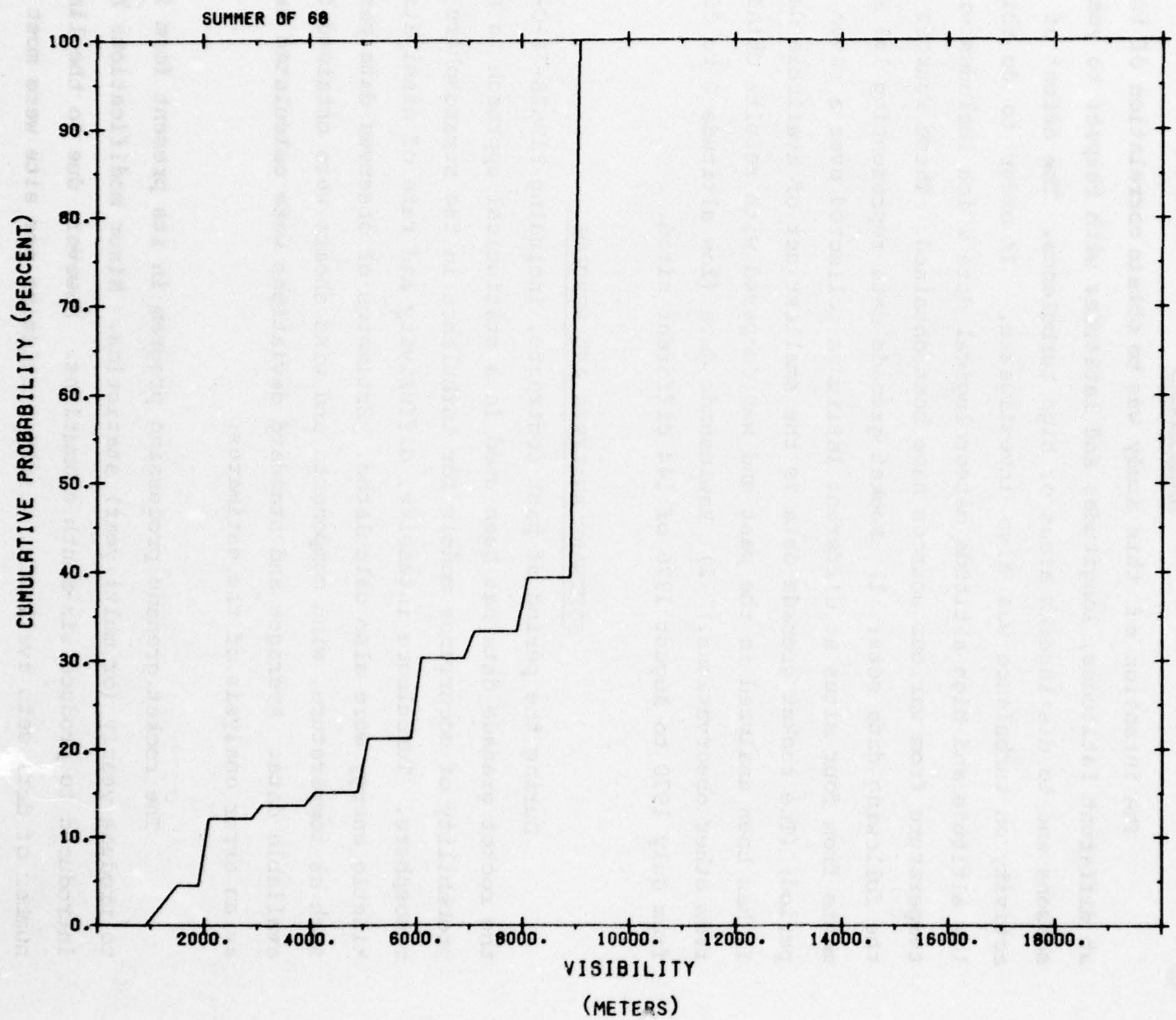


FIGURE 2-4 CUMULATIVE PROBABILITY OF VISIBILITY

3. TURBULENCE ANALYSIS

INITIATOR: E. MURPHY PROJECT NO: 6687 PROBLEM NO: 4889

BACKGROUND

The intention of this study was to obtain correlation of turbulence at different latitudes, longitudes and latitudes with respect to yearly seasons and to distinguish areas of high turbulence. The effect of solar activity on turbulence was also investigated. In order to do this, low altitude and high altitude meteorological data which includes winds and temperature from various sources have been obtained. These sources include the following data sets: 1) Rocket grenade data representing 153 measurements from four sites at different latitudes collected over a seven year period; (The rocket grenade data is the smallest set of available data. It has been analyzed in the past and was compared with results obtained from other observations.) 2) Rawinsonde data (low altitude 0 to 35 km.) from July 1970 to August 1976 of 144 different sites.

SOFTWARE ANALYSIS AND RESULTS

During the period of past contracts, including F19628-78-C-0157, the rocket grenade data has been used in a statistical approach to provide probability of occurrence models for turbulence in the stratosphere and mesosphere. Turbulence intensity, diffusivity and rate of dissipation of kinetic energy were also calculated. Estimates of observed data parameters, such as temperature, wind components and wind shears were obtained for all available data. Averages and standard deviations were calculated as well as an error analysis of the estimates.

The rocket grenade processing program in its present form is used to produce yearly (or multi-year) statistics. Minor modifications have been introduced to produce six-month summations. However, due to the limited number of data sets available, yearly statistics per site were most often computed.

The basic input for this program is temperature and wind component values per altitude per experiment. The program first initializes the arrays and counters and then reads the data set. If the entire data set has been read, the means and standard deviations are computed. If the entire data set has not been read, the program checks for valid data, and if the data is valid, performs a spline interpolation and sums the necessary parameters. If the data is not valid, the set is ignored and a new data set is read. Once the means and deviations have been computed, statistics are listed. The program then determines which layers, if any, require further summation for the averages of the averages. When all summations have been completed, this output is produced.

The following equation is used to determine the Richardson number (R_i). The Richardson number is the basis of the criterion for the presence or absence of turbulence.

$$R_i = g \left(\frac{\partial T}{\partial z} + \Gamma \right) \Big/ T \left(\frac{\partial u}{\partial z} \right)^2$$

where g = the gravitational value at the latitude - assumed to have negligible error

Γ = dry adiabatic lapse rate which is a generally accepted constant (9.8°k/km)

Turbulence is considered present when $-4 \leq R_i \leq 0.25$. Some analysis has also been performed for $-4 \leq R_i \leq 1.00$. The remaining equations for the calculations of the turbulence parameters can be found in the final report, Mathematical Analysis and Implementing Software for Physical and Engineering Data, 17 May 1977 - 16 May 1978.

There are three basic outputs produced by the rocket grenade processing program with an option for a fourth. Output number one produces per altitude: number of sample, number and percent turbulence occurrence, means and standard deviations of Richardson number, rate of dissipation of kinetic energy, turbulent intensity and turbulent diffusivity. Output number two produces per altitude: means and standard deviations of the spline interpolated input parameters, (temperature, east-west and north-south components.) Output number three produces, for input specified layers, averages of the averages produced for output number one. Output four, if selected, produces per altitude: temperature, wind shear and turbulence information for every interpolated input point being processed. Outputs one through three are always obtained. Output four is only needed when anomalies appear in the data set and the location of the anomaly must be found so that the particular data set may be ignored.

During the period of the past contract, modifications were made in this program. One alteration modified the calculation of turbulence parameters. The rate of dissipated kinetic energy and turbulence diffusivity calculations are now based on the minimum value of three calculated lengths of turbulence. The program was also adapted to sum the number of turbulence occurrences for $0.25 < R_i \leq 1.00$ so that an analysis using an expanded Richardson number interval of $-4 \leq R_i \leq 1.00$ can be performed, if desired.

The program generated to process the rocket grenade data was modified to process low altitude data. The four basic differences between the low altitude program and the rocket grenade program are:

1) The low altitude program performs a seasonal analysis, where as the rocket grenade program performs a yearly or six-month analysis;
2) The rocket grenade program performs an analysis on high altitude data (35 to 95 KM) as opposed to the low altitude range of 0 to 36 KM;
3) The averages of the averages computed in the rocket grenade program are not computed in the low altitude seasonal program (they are calculated in a separate program). The amount of low altitude data is orders of magnitude greater, hence, a greater and more comprehensive analysis could be accomplished.
4) During the period of the contract, changes were made to calculate the statistics at .1 KM increments as opposed to the original 1.0 KM increments.

The data sources for the two programs are significantly different. The rocket grenade data was available on cards and presented no data handling problems. However, the rawinsonde data was sent via tapes and the data extraction was the first task to be completed.

A program was written to extract low altitude rawinsonde data from 9 track tapes. At present, 40 tapes containing approximately 150 stations, and 7 years of data (from 1970 to 1976) are available. Data is available for the lower altitude (0 - 35 Kilometers). The program searches an original rawinsonde data tape for a particular station and year. Program LOWALT then begins data extraction by buffering in a variable length of blocks from a 9-track magnetic tape. ... set of blocks does not exceed 6000 bytes and all data are binary.

The first decode in the program unpacks the first 33 bytes. The following information is obtained: the block length (this occurs once at the beginning of each block), deck number, station number, year, month, day, hour, and number of levels. The second decode unpacks 25 bytes of information. The first level at the beginning of the block is the first level of altitude, generally the surface level. The format allows up to 79 levels. All levels contain pressure, height, temperature, humidity, wind direction and speed, height indicator and temperature indicator. All levels have the same length (25 bytes). The third decode is used to check for the last record in the block. If the information corresponds to the first set of data, then this block is done, and the program will go on to read in another block of data. The program will print the station, year and month, and write the extracted data on an output multi-file tape, each file containing data for a specific site and a specific year.

Although problems were encountered with this task at first (refer to final report, Mathematical Analysis and Implementing Software for Physical and Engineering Data, 17 May 1977 - 16 May 1978), the format of the tapes available during this period were the same as previously preprocessed tapes. Approximately 600 site-year combinations were preprocessed.

The new multi-file tapes generated by the program LOWALT become the basis of a tape library which saves the original data for future analysis. There are currently over 200 tapes of this type in the library, each containing approximately seven different site-year combinations.

The second generation multi-file tapes are then used for a seasonal analysis. The program in its present form is used to produce seasonal statistics per station for one year's data. The basic input for this program is temperature and wind component values per altitude per experiment. The basic flow of the low altitude seasonal program is like that of the rocket grenade program, the major difference being that after the initialization of the arrays and counters, it is necessary to read data from the last ten days of the previous year, if such data exists, in order to have a complete data set for the winter season of the year being processed. The program also writes onto a permanent file the data for the last ten days of the year being processed, so that it may be used when the next year's data is processed.

There are two basic outputs, with options for a third and fourth. Output number one produces per altitude: number of samples, number and percent of turbulence occurrence, means and standard deviations of the Richardson number, rate of dissipation of kinetic energy, turbulent intensity, and turbulent diffusivity. Output number two produces per altitude: means and standard deviations of the spline interpolated input parameters (temperature, east-west and north-south components.) Output number three, if selected, produces per altitude, the input parameters: temperature, north-south wind speed and east-west wind speed. Output four, if selected, produces per altitude: temperature, wind shear, and turbulence information for every interpolated input point being processed. Outputs one and two are the basic outputs and are the only two normally obtained. Outputs three and four are useful when checking an algorithm change or if anomalous data exists.

The statistics generated are saved on a new multi-file tape with each file containing data for a specific site and a specific year. These tapes of generated seasonal statistics become part of the tape library mentioned above. There are presently six multi-file tapes containing statistical data, each with approximately 50 different site-year combinations.

The third task in this sequence reads the statistics generated above and averages them over intervals specified by the user. The purpose of this task is to aid in the determination of the significant altitude levels at which pronounced occurrences of turbulence are found in the troposphere and how their levels vary with season and latitude.

This program, written during the period of the past contract, will read seasonal data sets and generate statistics for any number of specified intervals. The basic input per altitude is: the number of samples, the number and percent of turbulence occurrences, the means and standard deviations of the Richardson number, the rate of dissipation of kinetic energy, turbulent intensity, turbulent diffusivity, and also the means and standard deviations of the spline interpolated input parameters (temperature, east-west, and north-south components.) These values are then averaged over the specified layers. Layer summations are performed exactly the same as in the rocket grenade program.

The last program written for this study displays the values of percent of turbulence occurrence. The program is an adaptation of an existing plotting package and uses the output generated by the averaging program. For each interval, it produces a seasonal plot of the percent occurrence for either $-4 \leq R_i \leq 0.25$ or $-4 \leq R_i \leq 1.00$, as specified in the input. The values displayed are done so at the proper site, latitude and longitude. The program is currently set up to do only one year at a time.

The program first reads the plotting limits for the x and y axes and which percent of occurrence will be plotted from input. A data card is then read, and a check is made to see if data was available for that season, altitude level and year. A new card is read if the sample was zero, otherwise, the data is sorted according to seasonal level. The latitude, longitude and percent of occurrence are saved in an array. An index is incremented for each season and level combination to determine the number of values to plot. If one of the index values exceeds 200, the array is exceeded and the program will stop and print a diagnostic message.

Once all the input cards have been processed, a do loop is set up to plot each set of latitude, longitude, and percent of occurrence combination, for the different season and level matrices. There will be four seasonal plots for each different altitude layer. The program is currently set up for four layers, therefore 16 plots being produced per run. Examples are shown in Figures 3-1 and 3-2. The most recent effort produced plots for 40 different stations for the years 1970-1976.

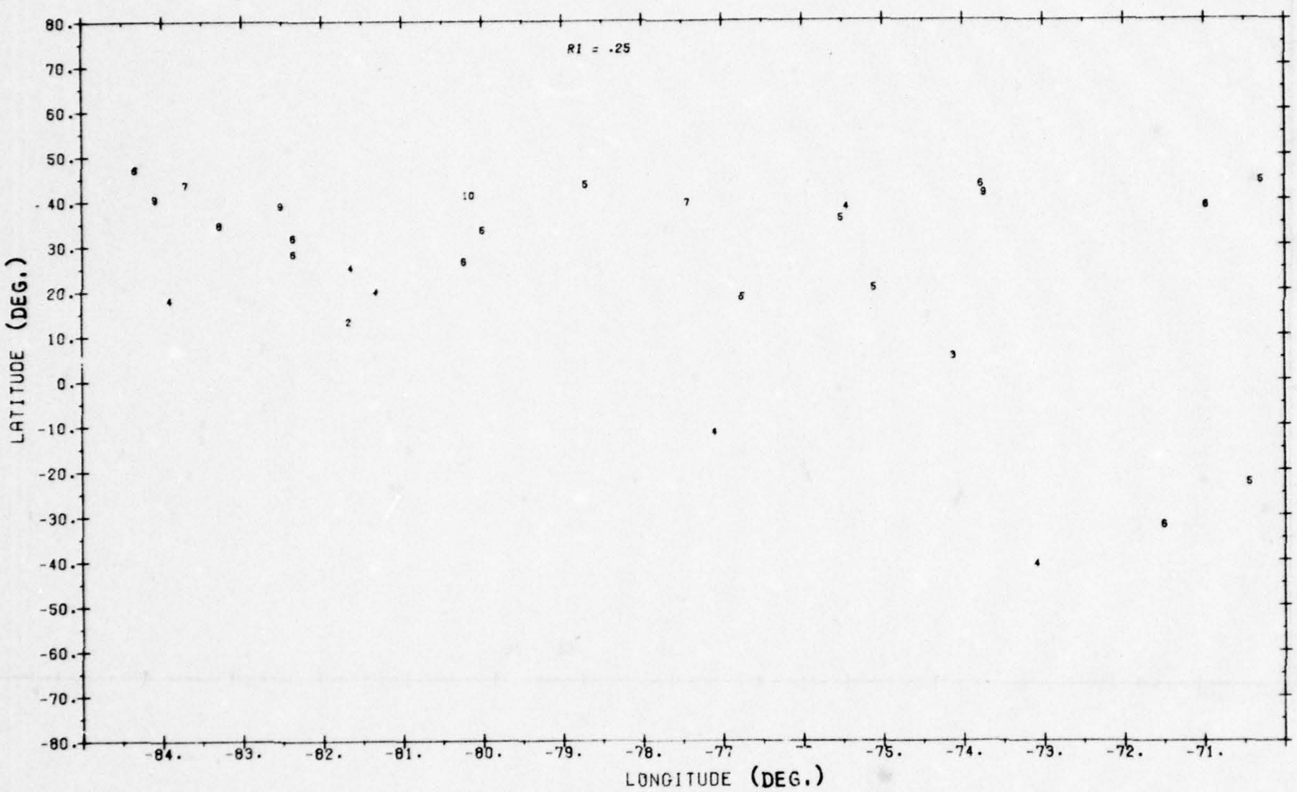


FIGURE 3-1 PERCENT OCCURRENCES FOR SPRING 1974 AT 0-5KM

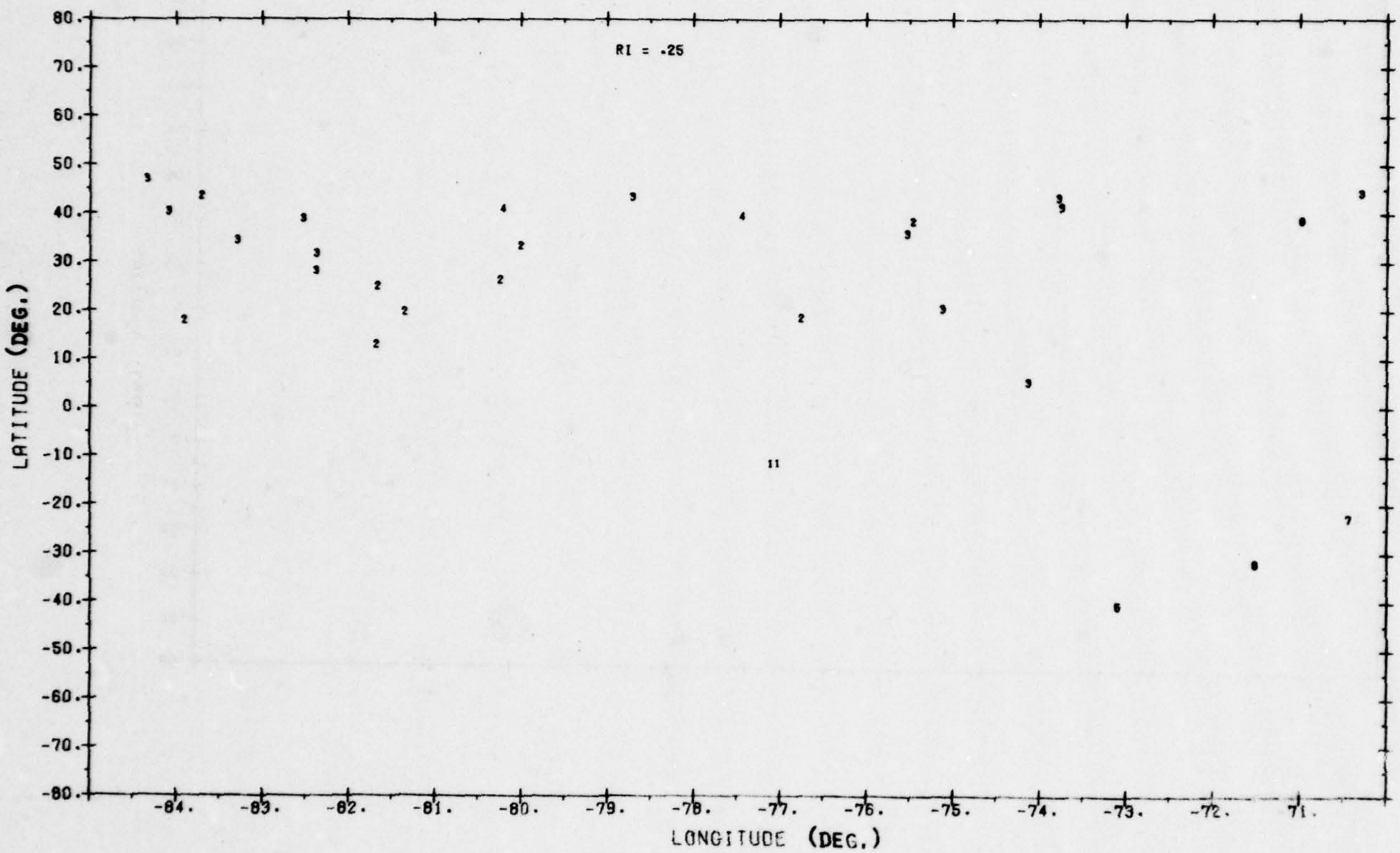


FIGURE 3-2 PERCENT OCCURRENCES FOR SPRING 1974 AT 5-10MM

4. SSJ/3 SENSOR DATA OBTAINED FROM DMSP SATELLITE

INITIATOR: P. ROTHWELL PROJECT NO: 2311 PROBLEM NO: 4961

BACKGROUND

The purpose of this problem was to reduce and analyze the SSJ/3 data obtained from the DMSP satellites. The SSJ/3 high energy electron data was extracted, printed and plotted. SSJ/3 data was obtained from the tapes sent by AFGWC and NOAA.

The SSJ/3 sensor is designed to detect and analyze precipitating electrons in the energy range from 50 eV to 20 keV. It is one of the supplementary sensors on board the Block 5D configuration satellites of the Defense Meteorological Satellite Program.

General descriptions of the structures of the AFGWC and NOAA tapes appear in sections 4 and 6 of the final report of May 17, 1977 - May 16, 1978, Mathematical Analysis and Implementing Software for Physical and Engineering Data.

SOFTWARE AND ANALYSIS UPDATE

In the last four months, most of the work involved in this task has centered around the production of plots displaying the functions of SSJ/3 data as function of time; the three parameters of interest are the average energy, energy flux and integral flux; and can be derived with the following equations:

1) average energy: $$\frac{\int_{1}^{16} E_j(E)\,dE}{\int_{1}^{16} j(E)\,dE} = \frac{\sum_{i=1}^{16} E_j(E)\,\Delta E_i}{\sum_{i=1}^{16} j(E)\,\Delta E_i}$$

2) **energy flux** : $\int_{1}^{16} E_j (E)\, dE = \sum_{i=1}^{16} E_j (E)\, \Delta E_i$

3) **integral flux:** $\int_{1}^{16} j (E)\, dE = \sum_{i=1}^{16} j (E)\, \Delta E_i$

where

$j(E)$ = normalized SSJ/3 data at particular energy level
ΔE_i = change in energy

Initially, each plot corresponded to 15 minutes worth of data. The area of interest was where the absolute value of the magnetic latitude was above 50°. Sample plots were generated and labelled with the corresponding ephemeris.

This software which calculates, integrates and plots the functions of SSJ/3 data as function of time was modified to produce 25-minute plots instead of 15-minute plots. 25-minute plots were better representations of the orbits than 15-minute ones. Two 15-minute plots were needed to display an entire orbit.

The only data not used in the 25-minute plots was where the satellite was over the area between -48° and 48° (magnetic latitude). 25-minute plots were generated for September 18, 19 and December 12 using both the AFGWC DMSP and NOAA data tapes. These plots were compared from both sources of data to determine which source would be used in future analysis. Upon observation of these plots, the NOAA tapes were picked to be the main source of the SSJ/3 data since these tapes were more complete than the DMSP tapes. Further, each NOAA tape contains approximately 9½ days worth of data while each DMSP tape contains dawn orbits of one day and the dusk orbits of other.

It was requested to produce 25-minute plots for all available data that was not of the restricted area starting in the month of December continuing to the present date. Initially, CRT plots displaying 25 successive minutes of data were generated for the month of December. To facilitate storage, software was then written to produce 25-minute plots on microfiche. Work was hampered because of numerous hardware problems. It was requested that when the microfiche was completely operational, all 25-minute data plots be done on microfiche instead of 35 mm CRT film. All available December and January data have been plotted on microfiche.

Modifications to the 25-minute plots were done to make the plots more readable (including enlarging the labels on the plots and space between each line of the label). In addition to these cosmetic modifications, further changes to the processing software was necessitated. These changes include:

1) The software used to extract the SSJ/3 data from both the AFGWC DMSP and NOAA data tapes was modified due to the changeover in the system. The previous extraction programs were not operational due to two bit and character routines.

2) Further, we were informed that the geomagnetic latitude, longitude, and time on the NOAA tapes were incorrect. A subroutine was obtained to calculate the correct geomagnetic coordinates. It was decided that the received subroutine CGLALO would be added to the extraction software to obtain the correct geomagnetic coordinates. Two additional values had to be extracted from the NOAA tape in order to use this subroutine; i.e. corrected geographic coordinates. Modifications to the extraction program have been added to obtain extra values.

FUNCTIONAL DESCRIPTION OF SOFTWARE

There are three steps in the production of 25-minute plots on microfiche. Step 1 involved the extraction of data from the NOAA tapes plus listing of actual time periods. Step 2 generated a new tape with a correct time sequence in order to execute step 3, the actual plotting of data. In step 1, each day of the month was extracted from NOAA data tapes and only the necessary data was stored on a multifile 9-track tape. Data is extracted from the NOAA tape with the aid of the subroutine BFF. Subroutine BFF reads 560 36-bit words into CDC 6600 60-bit words and unpacks each 36-bit words into 60-bit left-justified words. Program handles data blocks of 20 words at a time where BFF results in 560 60-bit words in each call to it. Location blocks plus four seconds of sensor data are unpacked upon determination of whether the data set contains actual SSJ/3 data or the information block. The SSJ/3 sensor data was shifted onto the NOAA tape in one set of 144 bits, re: 16 channels occupying 9 bits per channel which makes up 4 words. Unpacking of the data words results in logarithm representation of the actual counts. Mantissa is represented in the five least significant bits while the exponent is in the remaining four bits. The actual count represented by the 9-bit word is given by:

$$COUNT = 2^{Y} (X-32) - 33.$$

Only data whose absolute value of magnetic latitude is above 45° is written on the storage tape. A call to the subroutine CGLALO is necessary to determine the magnetic coordinates. The subroutine CGLALO is the Hakura/Gustufsson coordinate transformation of geographic latitude and longitude to corrected geomagnetic latitude and longitude.

Procedure continues with calls to the subroutine BFF when more data is needed until a data set is encountered whose Julian date is day after day of interest. At this point, the storage tape is rewound and the actual time periods for this day are listed. The program keeps track of the time and when there is discontinuity, the time of the first and last data sets in the group before discontinuity are printed. The time listing illustrates which data sets are out of time sequence or duplicated also determines the completeness of a days worth of data and where bad portions of the data exist.

In step 2, the program reads in one data card stating how many groups of data sets are out of sequence, initial and final times of each groups, and expected locations of the groups to be placed. Only one or two groups can be placed in their destinated areas; modifications can be added to the existing program to handle additional out-of-sequence data sets. The program has ability to eliminate individual duplicate data sets; it cannot handle a large group of duplicated data sets. A new tape is generated with a correct time sequence inserting the one or two misplaced groups of data at their destinated areas.

Actual time periods on the new tape are listed. To save storage space, most of these storage tapes were 9-track tapes consisting of 5 or 6 days of data. Earlier runs resulted in 7 track tapes with three days of data. Data stored on the 7 track tapes have been copied over to 9-track tapes.

In step 3, input tape generated by step 2 is read to obtain four seconds of sensor data along with the ephemeris data. The data set is checked to make sure it is in the desired area; i.e. absolute value of magnetic latitude above 48°. Subroutine CGLALO is necessary to obtain magnetic coordinates. Problems have been encountered with the time on the tape, therefore a few checks on the time are included in the program; e.g. seconds being greater than 60. There is an option to start the program at a destinated time by two simple changes; re: setting CHECK to 1 and changing conditional "if" statement.

Satellite revolution (rev.) numbers, used as a label on the plots, were obtained from the information blocks of the NOAA tapes. A permfile was generated containing the rev. numbers along with their destinated days and times. The plotting program will search through the permfile until a rev. number is found for specific day and time or an end-of-file is encountered. If there is no rev. number for this time of day, rev. is set to -99.9, indicating missing data.

Each plot has at most 25 successive minutes of data. When the first possible point to be plotted is determined, the bounds are set for both X- and Y- axes. On the plot, tic marks appears at every minute of 25 minutes.

When the bounds are set, the times in seconds at each tic marks are calculated. Geographic coordinates, magnetic coordinates, and magnet. local time must be determined at each minute and stored for the plotting routine. There is some overlapping of data sets on the tape; to eliminate this extra data, a flag system is present in the program. The flag system looks at one hour at a time. When a new hour of data is encountered, the flag system is zeroed out.

Each data set contains four seconds of sensor data but each second is reviewed individually. Equivalent flux at each of the 16 channels must be determined to obtain the functions of SSJ/3 sensor data. Dividing the count rate by the approximate normalization factor gives the equivalent flux in electrons per cm^2/sr/er/sec. Normalization factors appear in Table 6-1 of the final report of May 17, 1977 - May 16, 1978, Mathematical Analysis and Implementing Software For Physical and Engineering Data. Since the equivalent flux values of channels 8 and 9 are almost equal, the average of the two values is used in the computations instead of the individual values of the equivalent flux.

The three functions at each second are stored in arrays VAL1, VAL2, and VAL3 along with time, GMT. To achieve the proper scaling, the three functions are divided by a factor of 0.098. If either energy flux or integral flux is zero, the three functions are not stored in the arrays for that second.

The procedure continues until 25 successive minutes of data are accumulated or a data set is encountered that was out of interested area; i.e. absolute value of magnetic latitude was below 48°. If the path has not been completely represented in these 25 minutes of data, i.e. there is more data following these 25 minutes in the same orbit, a call to the subroutine ADJUST is done. The subroutine ADJUST eliminates any data stored in the arrays for the first minute and adds onto the arrays, the next minute of data following the 25 minutes period. The bounds are adjusted to reflect this change as well as the storage arrays of ephemeris. A call to the subroutine FCTN is necessary to generate three plots illustrating average energy, energy flux and integral flux as functions of time. At every other tic mark beneath the plot of integral flux vs. time, the following is listed: time in seconds, geographic and magnetic coordinates, and magnetic local time; the values for time and geographic coordinates were obtained from the tapes while magnetic coordinates were calculated in the subroutine CGLALO. All but the magnetic local time were passed to the subroutine FCTN. The magnetic local time is calculated in the subroutine FCTN using the following equation:

$$MLT = UT + \frac{MLON}{15} + 18.67$$

where UT = Universal time
MLON - Magnetic longitude

If MLT is greater than or equal to 48, 48 is subtracted from MLT, while in case MLT is above 24 but below 48, the value of 24 is subtracted from MLT.

Each value to be listed is passed to the subroutine MORE where the proper spacing is determined for the labels. The functions along with the corresponding ephemeris and number of points plotted are written onto a 9-track tape multifile storage tape for future analysis.

ADDITIONAL SOFTWARE

Software was written to produce CRT plot of equivalent flux as function of center energy for destinated time periods. Each plot contains 10 spectra for 10 successive seconds of data. The data was plotted on a log/log scale and the plots were labelled with the ephemeris data pertaining to the first second of 10 second set. Program generated a destinated number of successive plots in the area where the absolute value of magnetic latitude was above 50^{o}. Sample data was processed.

The program that generates CRT plots of equivalent flux as function of center energy must be altered to accept a new tape generated by existing NOAA extraction program because two additional words appear in each record. Records have been written on tape with binary format.

Each record of sample data contained 4 seconds of sensor data plus ephemeris. The data set is checked to make sure it is in the interested area; i.e. absolute value of magnetic latitude is above 50^{o}. A call to the subroutine MAG is necessary to obtain the magnetic coordinates. The magnetic coordinates are calculated using the following equations:

$$lat_{mag} = 90^{o} - \text{arc cos}(\sin 78.56 \sin (lat_{geo}) + \cos 78.56 \cos (lat_{geo}) \cos(long - 69.76))$$

$$LONG_{m} = \text{arc sin} \left[\frac{\sin (LAT_{c}) \sin (LONG - 290.24)}{\sin (90 - LAT_{m})} \right]$$

where lat_{mag} = magnetic latitude; $long_{m}$ = magnetic longitude;
lat_{geo} = geographic latitude; long = goegraphic longitude;
lat_{c} = colatitude = $90 - lat_{geo}$

Modifications should be made in the calculations of the magnetic coordinates by replacing the subroutine MAG with the subroutine CGLALO. The subroutine CGLALO gives a better approximation of the magnetic coordinates.

Equivalent flux at each of 16 channels is calculated at every second of interested period of time by dividing the count rate by the approximate normalization factor. All non-zero equivalent flux values are divided by 0.098 to allow proper scaling and are stored in array Y.

On each plot, ten spectra are displayed for 10 successive seconds. When the first spectra is plotted, the following is labelled on the plot: Julian date; starting time in hours, minutes, and seconds; magnetic coordinates; which has been passed to the subroutine SMILOG. For each new spectra, the origin is redefined (0.5, 0.5) from previous location. When the tenth spectra of the set is plotted, a new plot is set up and the counter is increased by one. This program will stop when the counter is the value of 5.

5. STATISTICAL ANALYSIS OF HIGH ALTITUDE MAGNETIC AND PARTICLE DENSITY DATA

INITIATOR: H. GARRETT, PROJECT NO: 7661, PROBLEM NO: 4935

BACKGROUND

Software necessary to analyze data from the ion and electron spectrometers on board the ATS-5 and ATS-6 Satellites was written by ASEC. Software to unpack and to convert this data was written and new preprocessed tapes generated. In addition, bad data points were eliminated and files containing data averaged over each 10 minute period was created. A plotting routine has been generated which plotted the individual energy density, energy flux, number density, and number flux for each ion and electron spectrometers (the parallel and perpendicular components plotted on one plot) on separate plots, each 8" by 10". Once the averaged, "cleaned up" data set has been generated, various programs were written to generate statistics necessary as input to a substorm environmental model required for spacecraft charging standard and design guidelines.

DATA OBTAINED AND SOFTWARE DEVELOPED

DATA CONVERSION

Three sets of tapes, two containing ATS-6 data, one containing ATS-5 data, were received. The first set of data tapes containing ion and electron spectrometer recordings generated on board the ATS-6 satellite were written on a UNIVAC 1108 36-bit machine. A source deck of the program that accesses the tape and provides listings of the data was obtained. The program was developed for use with CDC 3600 computer. In order to obtain listings, the input/output sections of the program had to be modified to be compatible with the CDC-6600 60-bit word. After conversion of the data, the program then converts the 10-bit log compressed data into particle counting rates, obtains distribution functions, prints the counting rates and dF for the four detectors and plots the results. (Refer to the documentation for the UCSD ATS-F Data Reduction Program for additional information on the original program.)

The program received expects data in the form of 48-bits of information per word. Since the CDC-6600 reads 60-bits of information per word, it was necessary to insert a subroutine in the program which unpacks the 60-bits of information into 48-bits right justified with zero fill, thus making it compatible with the rest of the input. The first record of each file contains characters, which must be decoded during initial processing. Because the BUFFER IN can be in either binary or character mode, but not both, a conversion had to be written up to convert from external BCD, to internal BCD. This was also included as a subroutine. Other minor modifications were necessary for tape access and positioning, logical If statements, and other minor compiler differences. Once these changes were made, the only modification left was in the output array. Instead of placing the output in array LEFLIN, the output went directly to the printer. In addition, routines were added to plot the differential function versus energy levels.

Changes in the logical and mathematical portion of the program were not made. To obtain the plots, 2 new arrays were set up to accumulate the energy levels and the differential functions and pass these to subroutine to label the plots. Refer to the UCSD ATS-F Data Reduction program for an explanation of the mathematical and logical procedures.

With the second set of data tapes containing information from the ion and electron spectrometer on board the ATS-5 Satellite obtained, it was requested that these tapes be converted to compatible CDC-6600 SCOPE tapes and the data listed. Program UPACK was written to do this and to obtain, in addition, the theoretical distribution which was also saved on the tapes. Multi-file tapes were set up for data storage and later used for plotting and statistical analysis.

This program buffers in one block of data, 300 UNIVAC 1108 words at a time. Since the UNIVAC 1108 contains 36-bit/word, this means that 3 CDC words contain 5 UNIVAC 1108 words. These values are all integer which helped to simplify the conversion. This was accomplished by setting up a do loop and decode 5 words in each group by shifting 36 bits in succession into one CDC 60-bit word right justified. The program then prints the pertinent information and writes it onto an output tape.

This program also has the option to calculate a theoretical differential energy flux (DEF). The DEF is calculated using the count rates and energy level extracted from the ATS-5 Satellite data. All data after conversion is written on a SCOPE standard tape for future processing.

The equations for the Maxwellian distribution and four of its moments are approximated within the program. These moments have been derived from the University of California San Diego ATS-5 data and can be used to derive a Maxwellian and a "2-Maxwellian" fit to the distribution function.

Program UPACK first initializes parameters then calls subroutine BFF which BUFFER'S IN data and converts each UNIVAC 1108 word into a CDC-6600 word by use of the ENCODE/DECODE function. One record is processed at a time. A check is made on the first and second word to determine if it is a header record or a data record. If it is a header record, the day, month, year are printed, and the energy levels are extracted and saved. A new record is then processed. If the record is a data record, the time is saved and printed, along with the rest of the parameters stated in the output section. The energy levels are converted to electron volts. This program via subroutine DIFF has the option to calculate from the electron and ion count rates (C_1, C_2) and energy (E), to calculate the differential energy flux, to obtain DF1, DF2 which are distribution functions for electrons and ions, and D1, D2 which are the approximation distribution functions.

The Maxwellian distribution function as well as number density, number flux, pressure and energy flux were calculated. The Maxwellian distribution function is of prime importance in various techniques used to calculate spacecraft potential.

The third set of data tapes containing ion and electron information recorded on spectrometers from the ATS-6 Satellite were processed, listed and stored on NOS/BE tapes for future use. Data along with the ATS-5 data were used as input to a substorm environmental model required for spacecraft charging standard and design guidelines. The data saved includes date, time, temperature correction coefficient, angle the N-S detector makes with the spacecraft, pitch angle, ion and electron number density, ion and electron energy density, ion and electron particle flux, ion and electron energy flux, and potential of the spacecraft.

For the third set of data tapes, program RDTPE first sets jl = 500 to prevent an end-of-file from being written at the beginning of the output tapes. The first block of data is then read by use of the FORTRAN BUFFER IN. This block contains 105 words. A check is made for END-OF-FILE. If there is a data record, processing continues even if there is a parity error. If there is an end-of-file, the output tape is rewound, the first record printed, and the program stops.

After a record is brought into core memory, the first 263 characters of the day are checked. If the day is greater than the previous record, an end-of-file is written on the output tape. The first record is then written on the output tape and printed. A do loop is then set up to process the remaining three records of the block. The characters of the record previously processed are skipped and the next 263 are decoded. Once the four records are processed, more data is brought into core memory, and the process continues until an end-of-file is reached.

AVERAGING

Output generated from data sets two and three were further averaged and "cleaned up". Known erroneous data values were eliminated. Data was checked with spectograms to determine the validity of the values. Programs

MNT1 and MNT2, were written to read the tapes generated by the ATS-6 and ATS-5 data tape conversion programs. Ten minute averages are obtained from the data extracted from these tapes, and bad points are discarded. The data is plotted and a permanent file is created containing the averaged data to be used for future analysis. Program MNT1 has the option of plotting the calculated distribution function as well as the theoretical distribution function. These programs will process up to the last day specified from input, and can either generate the permanent file and/or plot the results.

These programs are part of a package used to generate input to a substorm environmental model required for spacecraft charging standards and design guidelines. These programs first read an input card which selects the plot and tape generation options, plus the beginning and ending days for processing the data. All parameters are then initialized. The input tape is read. All records are checked for end-of-file and parity. If there is a parity error, a new record is read. The parameters that need smoothing are saved in an array and another record is read.

The other parameters are summed by checking the input time with a preset time of 10 minutes. When this time is exceeded, the sum is divided by the total number of points and this value saved. Parameters are initialized and the processing continues.

When an end-of-file is reached, or when a new day of data is read, the data stored in memory is smoothed in subroutine REMOVE.

The standard deviation for the entire set is calculated using the equation,

$$STD = \sqrt{\frac{\Sigma(X-\overline{X})^2}{N-1}}$$

which is used in the smoothing algorithm.

A 20 minute running average about each value is calculated and if the difference between this average and the data value exceeds twice the standard deviation, the value is skipped. The time and data value are placed in new arrays and returned to the calling program after all values are checked.

The smoothed data is then plotted if this option is selected. The data is averaged in increments of ten minutes and the average data is placed in a 2-dimensional array (150, 8) for MNT2 and (150,16) for MNT1 to facilitate plotting. A do loop is set up to plot the values and before each call, a label identification is written.

Before the average data is plotted, this data along with the other parameters are placed in an output array (150 values maximum) and written on tape. One record being equal to one 10-minute average data set. This process is repeated for each day.

For the ATS-5 Data set, (program MNT1) the energy levels and distribution functions are read separately, and saved for plotting. Label information is set up in the main program and the values are plotted in subroutine PLT. Different header subroutines are used for the electron and ion differential distributions. Parameter IPLOT determines how many of these plots can be generated. Either a log-log and/or log-linear plots of these parameters can be generated.

INCORRECT DATA REMOVAL

A program was written to remove known incorrect or anomalous data values. This program reads in from data cards the times of the anomalies for the ATS-5 and ATS-6 data sets. The program checks through the permanent file which contains the time of the faulty data until it finds the proper record. The program then will replace the defective value(s) with a value of -999999.0. This predetermined value will be a signal to future

THIS PAGE IS BEST QUALITY PRACTICABLE
FROM COPY FURNISHED TO DDC

PLOTTING SOFTWARE

To display the extracted ATS-5 and ATS-6 data, software was written. Program DFCTP5 and DFCTPT reads ten minute averaged data from a permanent file, and plots the number density, number flux, energy density and energy flux with respect to universal time for each day. These plots are arranged so that the four plots for the electrons and the four for the ions are side by side for comparison.

Program DFCTP5 processes the data from the ATS-5 satellite which contains the parallel and perpendicular components of the electron and ion detectors, hence 16 curves are plotted, two for each component/detector combination. Program DFCTPT processes the data from the ATS-6 satellite which contain only the perpendicular components of the component/detector combination. The input is slightly different for each case.

The programs read an input card which contains two parameters. The first controls whether plots will be from CALCOMP or from CRT. The second controls the last day which will be plotted. Parameters and arrays are then initialized, and the first record of the data file is brought into memory. The date and time is printed out, and the parameters are multiplied by factors which enables the ATS-5 and ATS-6 satellite data to be easily compared. A header label is placed on the plot for each group, and by use of a do loop, 16 or 8 plots are plotted. A check is made for the last day to be processed. If it is not, the above procedure continues until the last day is reached, at which point the plot and programs are terminated and date and time are printed out.

Next by using a do loop and modulo 4, the factors are used for the ATS-5 and ATS-6 data to produce identical units. The factors for the ATS-5 and ATS-6 data sets are as follows:

Variable Description		ATS-5 Factor	ATS-6 Factor
Number Density	(No./cm^3)	10^{-3}	1
Number Flux	(No./cm^2-sec-ster)	10^4	$1/2\Pi$
Energy Density	(cV/cm^3)	.624	1
Energy Flux	(cV/cm^2-sec- ster)	6.24×10^7	$1/2\Pi$

When the last point of the day is processed, plot parameters are initialized and the plotting header subroutine STLAB is called. Header information is passed through COMMON statements from the main program, and values are ENCODED. The plot subroutine SYMBOL is called to plot the symbols desired from array BCD.

Program control now returns to the main program. The pen, or in the case of CRT plots, the beam, is positioned by the CALL to PLOT routine to prepare for the first plot. A do loop which goes from 1, to the maximum number of detectors, 4 for the ATS-5 data, and 2 for the ATS-6 data is begun.

The starting location of array Y1 is passed to subroutine SETPLT for processing. Parameter I1 is the 2nd dimension on the array. Axis label information is determined before the call to SETPLT.

In subroutine SETPLT, a do loop is set up to find the maximum value for each group to be plotted. The maximum and minimum value used for plotting are printed out. Another loop is set up to find missing values. If there are missing values for more than 1/2 hr., the previous values are plotted. This will cause gaps in the plots to show significant losses of data. If less than 1/2 hr. of data is missing, that data value is ignored and the plotting effect causes that value to be an average of the value before and after the missing value. This continues until all points are checked and plotted, then program control returns to main program. Subroutine SETPLT is called three more times for each of the moments before the loop is completed. When control passes beyond the do loop, a new frame of plot is advanced and a check to see if the last day has been reached. If it is not, parameters are re-initialized and processing continues. When the last day has been reached, the plot and program stop.

The Y-axis for each plot contains the same increments. Factors are printed as legends on the plot. These factors scale the plotted value to arrive at the real values. This was used to assure that all plots from different days would span the full range of the plot.

STATISTICS

Two packages of program were written to generate statistics. The first package of routines consists of several programs which will plot the histograms for temperatures, densities and spacecraft potential versus time. The main objective of the first program, TEMP, is to calculate the percentage of time that a geosynchronous satellite encounters substorms at various Electron and Ion Temperatures. The program first reads in the cell width size for both temperatures and then proceeds to calculate, if possible, the two temperatures for all days on the ATS-5 (or ATS-6) data set. Program TEMP will output two things. First, a listing of all the temperatures, number of valid temperatures, maximum temperature and minimum temperature. Second, histogram plots showing how the two temperatures are scattered over time. (Where data was fit by double or multiple Maxwellian distributions, both methods of averaging the temperatures were done).

The purpose of the second program, CDENS, is to show the percentage of time that a satellite encounters substorms at various Electron and Ion Current Densities. This program performs the same calculation as program TEMP, except for moments of the two densities instead of temperature are used.

programs that this data is to be skipped. After all records have been updated, a new permanent file(s) will be created.

This program first reads a data card which contains the time an anomaly occurred, the flags for the four detectors, their corresponding number density, number flux, energy density and energy flux. The program then reads the permanent files, that contain the ATS-5 or ATS-6 Data until the faulty record is found.

Once the record is found, the program then checks the flags to see which one(s) of the values are to be updated. The sixteen flags are in the same order as the sixteen moments on the permanent file. If the flag is equal to zero, the moment need not be changed. If the flag is equal to one, the moment is faulty and is to be replaced, by the value -999999.0. This predetermined value is a signal which indicated that the particular moment is faulty. The program then prints out a partial listing of the faulty record.

After all the anomalies are corrected, a new permanent file is created containing the updated records. The ATS-5 permanent files have an end-of-file marker at the end of each data set (approximately ten days). The ATS-6 permanent files have an end-of-file marker after each day on that file. Therefore, when an end-of-file is reached on the ATS-6 files, the program has to check if there are more days on that permanent file or whether it is to read the next permanent file.

The variable NOFS in the program stands for the number of permanent files that are being attached to the program for a particular run. When running the ATS-5 data set, NOFS can be any number between one and six. When running the ATS-6 data set, NOFS can vary between one and five.

The last program, POTENT, will show the percentage of time that a geosynchronous satellite is charged to various potential levels. The procedure is the same as the above two programs, except for the potential levels are plotted instead of temperatures or densities.

The first program, TEMP, reads in the initial cell size and minimum for both the electron and ion temperatures, then generates the remaining cells accordingly. The permanent files are then read, and the two temperatures for each record on all the files are calculated, if possible. If one or more of the moments used in calculating the temperature is missing (value of -999999.0), then the temperature can not be calculated. If all the moments are present, the temperature is then calculated by one of the Maxwellian distributions. The maximum and minimum values for each temperature, the counts in each cell and the total number of valid temperatures are then determined. When all the statistics for all the days on the permanent files are determined, a listing of these statistics are then output. When the list is completed, then the histograms for the two temperatures are plotted.

When the data was fit by double or multiple Maxwellian distributions, separate plots were requested for both methods of averaging, i.e.:

both $\frac{P}{N}$ and $\frac{EF}{NF}$ Where n = number density
p = pressure
NF = number flux
EF = energy flux

were calculated.

When processing the ATS-5 data set:

Electron Temperature (AVG) = average of parallel and perpendicular components times FACT 1

Electron Temperature (RMS) = average of parallel and perpendicular components time FACT 2

Ion Temperature (AVG) = average of parallel and perpendicular components times FACT 1

Ion Temperature (RMS) = average of parallel and perpendicular components times FACT 2 where FACT 1 = (6.24/(2/3))/10

where FACT 2 = 6.24/2

When processing the ATS-6 data set:

Electron Temperature (AVG) = average of perpendicular components times FACT 3

Electron Temperature (RMS) = average of perpendicular components times FACT 4

Ion Temperature (AVG) = average of perpendicular components times FACT 3

Ion Temperature (RMS) = average of perpendicular components times FACT 4

where FACT 3 = 2/3 and FACT 4 = 1/2

If one of the moments used in calculating either of the temperatures in the ATS-5 or ATS-6 data is less than zero, the temperature will be set equal to -999999.0 and then that temperature will be eliminated from any further statistics.

The second program, CDENS, uses the same procedures as TEMP, but instead of checking the moments used in calculating the temperatures, the ones used in determining the two current densities are checked instead. The same type of output also is given as the previous program but the electron and ion current densities are listed and plotted instead.

If one of the moments used in calculating either of the two current densities in the ATS-5 or ATS-6 data set is less than zero, the density is set of -999999.0 and eliminated from any further statistics.

The third program, POTENT, reads in the starting level for the space vehicle floating potential and then proceeds to determine how the values are distributed over these levels. The number of potential values in each level is determined and plotted. The maximum and minimum values are determined, the minimum non-zero value is found and the total number of potential values is listed.

This program reads in the potential values from the five permanent files and keeps track with two indicators, KOUNT and MZERO, the number of valid values and the number of times a potential value is greater than zero.

The number of potential values whose value is zero is included in the number of valid points, but not in the distributions. The distributions show how many non-zero values are in each cell and they are the values that are plotted against percent of time.

The second package consists of five programs which will calculate various cumulative matrices, means and standard deviations. These matrices consist of histograms such as temperature vs. density, temperature vs. time, temperature vs. KP values, potential vs. time and potential vs. KP values.

The main objective of the first program, TEMPDEM, is to display the ranges of current densities as a function of electron and ion temperatures. The program reads in the initial value and cell widths for electron temperature, ion temperature, electron current density and ion current density, then sets up the rest of the cells accordingly. The temperatures and densities are then calculated as well as the means and standard deviations for each 10 minute average on all the ATS-5 or ATS-6 permanent files. The program will output two 20 x 20 matrices, one electron temperature vs. electron current density, the other ion temperature vs. ion current density. The program will also output the means and standard deviations for all the temperature and density bins as well as for each cell in the matrix. (Where data was fit by double or multiple Maxwellian distributions, both methods of calculating the temperatures was done.)

The purpose of the second program, TIME, is to display the typical ranges of temperatures and current densities versus time in 3-hr intervals. The program reads in the cell widths for electron temperature, ion temperature, electron current density and ion current density, then sets up the remaining cells accordingly. The temperatures, densities and their 3-hr time interval (local time) in which they occur are calculated. After the time interval is determined the means and standard deviations are calculated for both the individual records on the permanent files and for each cell bin in the matrices. The program will generate as output four 8 x 20 matrices showing the two temperatures and the two densities plotted against 3-hr time intervals. The various means and standard deviations are also output.

The third program, DPTEMP, is set up the same as program TIME, but the difference is that the temperatures and densities will be plotted against the KP values rather than the 3-hr intervals. The means and standard deviations will also be different to reflect this change.

The fourth program, PTIME, uses the same procedures as program TIME, the only difference being that the floating potential values will replace the electron current density. The same output format will be generated as program TIME.

The fifth and final program, PKP, utilizes the same concept as program PTIME, but the potential values will be plotted against the KP values rather than the 3-hr time intervals.

The means for the individual cells and the cell bins were calculated in the following manner:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

The standard deviations were determined as follows:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

When using the ATS-5 data set:

Local Time = Universal Time -6.5

For the ATS-6 data set:

For year 1974, Local time = Universal time - (6 + 4/15)

For year 1976, Local time = Universal time - (21 + 2/3)

RECENT SUBTASKS

The most recent updates consisted of modifying these plotting routines (ATS-5 and ATS-6) to obtain plots such that on an 8" x 10" surface eight plots (each 2" x 4") are plotted. In addition, the data was multiplied by factors which displayed the ATS-5 and ATS-6 in the same units for comparison purposes.

This was accomplished by first reading the data from the input files and multiplying them by the appropriate factor The plot parameters were then set up, changes were made in the scale parameters to reposition the scales, and new scale labels created. The scales were made identical for all plots, and a factor which is printed on each plot was used to determine the proper value of each point. This factor is an integer value, and a point of the curve is obtained by multiplying the scale by the factor and using the power of 10 printed on the y-axis as the magnitude of the data.

The factor was obtained by dividing the expected delta scale value by the delta scale value used. Since all values used were rounded to the nearest whole number, integer values are obtained. Similar moments for the ATS-5 and ATS-6 data were plotted with the same magnitudes. For the ATS-5 data, the parallel component is a solid line, and the perpendicular component is a dashed line.

After a full set of these plots were generated, they were checked to determine their accuracy. There were many drop-outs and spikes in the data which were determined by the researcher to have been poorly created by the originator of the tape sent to him.

A new tape was received which contained the updated data sets to the ATS-6 data. This data was similar, but not identical to the previous tapes received. The data-items were not in the same order. The previous program was corrected and a new tape generated. The program which generates the 10 minute average data was run and a new file created. The old file was then merged with the new file by replacing the original data with the new data for those period which contained new data. Plots of the new ATS-6 data were generated and submitted by using the program described previously.

6. SOFTWARE SUPPORT FOR INVESTIGATING IONOSPHERIC SCINTILLATION DATA

INITIATOR: H. WHITNEY PROJECT NO: 4643 PROBLEM NO: 4893

BACKGROUND

Ionospheric scintillations data was analyzed by ASEC. The software for the analysis and extraction of this data was completed prior to the period of the contract F19628-78-C-0157. This software package included the unpacking of the digitized data tapes, obtaining the proper calibration information from the calibration data files, converting the data into dB values by using the calibration information, obtaining statistical analysis of the results for the full data sets in groups of one and a half minutes, and printing the results {this information includes the median of the data (in dB), the 98% value (dB), the mean of the power, S_4, Nakagami-m, the time for which the autocorrelation first is equal to or less than 0.5, the time at which the autocorrelation first equals 0.0, the deviation of these two values at this time, the maximum cross correlation and its associated time, the velocity of cross correlation defined as the recording stations separation distance divided by time of maximum cross correlation, and deviation of the cross correlation}, plots of the autocorrelations for 0 to 16 second time lags, plots of the cross correlation for 0 to 1 minute time lags, plots of the probability distribution functions and the cumulative distributions function, the calculation of the Chi-Square goodness of fit test, power spectrum, fade rate, and message reliability.

SOFTWARE UPDATE AND DATA PROCESSED

During the period of the past contract, six tapes were received for processing. Since the software had been completed, this required running the various programs to obtain the results. The operating system at AFGL was changed from SCOPE to NOS/BE at this time. This change affected the system software for creating multiple multivolume tapes. Since these were required in our present processing methods, changes had to be made to make this package more flexible.

The above problem was solved by running the program which unpacks the data along with the program which converts the data into dB values under one job. The only files saved in the unpacking process were the calibration files which are small compared to the data files and hence were placed on one tape volume. The remainder of the processing continued in the same manner as previously developed. The following is a description of the software as well as the functional steps required to process the data.

SCINTILLATION DATA ANALYSIS PACKAGE

The first step required in processing and analyzing ionospheric scintillation data available on a semi-regular basis from Peru involves the use of one of two programs, RTPE2 or RTPE4. RTPE4 is required when four channels of data are available and RTPE2 is used when only two channels of data are available on tape.

The programs convert data words from a tape written in Honeywell 316 format into CDC 6600 unpacked format, one data word per computer word. The digitized data words are written onto a NOS/BE standard 7-track tape, to be processed further by subsequent programs. Counts of the digitized levels in steps of 100, from 2100 to -2100 are printed out for the four channels after processing of each file. The only file processed at this point are the calibration files.

The program buffers in one block of data from the digitized tape and converts the first ten words in the buffer into CDC words which contain tape time and parameter information. These ten words are printed, for later use, and written onto a NOS/BE tape, for further processing. The remaining packed data words in the buffer are converted into individual CDC data words in the buffer are converted into individual CDC data words, and written onto tape. When an entire file has been processed, the file is reread, and the distribution of data is found and printed out, to assist in later calibration. This processing is continued until all files have been completed.

The digitized tape is attached as TAPE 10, and one block is buffered into the array JBUF. The first 2 words in JBUF are converted, using bit shifting subroutines, into 10 header words and written out. The remaining words are packed with five data words per CDC word. These words are unpacked, using bit-shifting routines, stored in the array IBUF, and also written onto tape. The number of words written in each record is given by the parameter ICHM which is printed at the beginning of the output, and as the ninth word in the ten word heading.

The distribution of data is found by counting the number of points falling into each of 41 cells, with values ranging from +2046 to -2047, in increments of 100. Values of +2047 and -2048 are considered as hardware zeros, and are counted in cell #42. The maximum and minimum values, as well as the distribution arrays, called COUNT and P, are set to zero at the beginning of each distribution count.

In the second step, the program averages data points from a calibration file, and prints out the points, the average value, and the count of the number of records read on the entire data tape, starting from the first file. This file is used to calibrate the data, and convert it into dB values. The numbers printed out correspond to voltage levels on the strip charts which are recorded simultaneously.

One calibration file is read in from the scintillation tape which was generated by either RTPE4 or RTPE2. This program can accept either 2 channel or 4 channel data. The file is read, one record at a time, the data is split up into the corresponding channels and the data values for each channel are averaged, 6 for a two channel tape, 3 for a four channel tape. Three parameters must be set to specify tape characteristics. They are at the beginning of the program. They are:

ICHS = number of channels on the tape.

IWD = number of words per record, corresponds to output parameter ICHM from program RTPE 4 or RTPE2.

IRCD = number of records between heading records, corresponds to output parameter NWR from program RTPE4 or RTPE2.

This average is printed along with the actual data values, the number of record, and heading words, which are useful for matching the calibration printout with the strip chart calibration. By equating steps on the strip chart with values on this printout, calibration levels can be found. The calibration ranges, along with the dB levels are then punched on a data card, to be used as input for the next program.

The third program in this task equates digitized counts of calibration signal levels to actual dB levels. The program first reads the input card which specifies the number of calibration levels and the channel to be processed. A check is made on value IFREQ to see if all cards are processed. A "999" is punched on this card to indicate all input has been processed. The input tape is rewound and then a do loop is set up going from 1 to the number of calibration levels as specified on the input card. The next input card is then read which gives the starting and ending record of the input tape for processing the first calibration level and the calibration value in dB. These values are printed out, and the input tape is read until the first record to process is reached. The calibration values are saved until the last record to process is reached then the average is calculated. The control returns to the beginning of the do loop and a new card is read. Processing continues until all levels for a channel are processed, then control returns to the beginning of the program where a new number of calibration levels and channel number are read, and the same process continues. The program is set up for four channel operation. Some calibration files have only one, two, or three valid calibrations. Uncalibrated channels are printed out as all zeros. After all channels are processed and a "999" is read from input the count number, averaged signal level and the corresponding dB level are printed out in a table.

From the table printed above, graphs of the calibration curves are drawn. These are then examined and possibly modified. When the graphs are approved, data is read from them and transferred to punch data cards, one card for each dB, containing the dB calibration and its corresponding scintillation value. There is one set of cards per channel. These cards are then used as input for the next task.

The next task involves the conversion of the digitized data into dB values. There are two programs available for converting the digitized data into dB values. The program WRITEDB for four channel tapes takes every four word, finds the two signal values which bracket it, and then finds the dB value corresponding to any intermediate value. A signal which is higher or lower than the maximum or minimum value allowed, is set to that limit. The tape is processed one record at a time, and the words per channel are written onto tape, so that there are a quarter as many words per record for each channel as there were on the input tape. All four channels are processed before another record is read in. At the end of the input tape, all four output tapes are rewound, and copied onto magnetic tape, one file per channel. The heading consists of only the four time words, which are used later in the analysis programs.

The above description applies to WRITEDB2 except that channels 3 and 4 have been eliminated. Only two channels of data are processed, half of the data words are channel 1 data, half are channel 2 data. Hence, for 2 channels, there are half as many words on the output tape as on the input tape for each data record. Program WRITEDB is used in conjunction with program RTPE. RTPE obtains the digitized values in CDC format and writes them on a dummy file. This file is then rewound and used for input to WRITEDB. These data files were too large to save for a single use.

When running either WRITEDB or WRITEDB2, the number of words per record and the number of records per block must be set at the beginning of the program. The arrays are large enough for 65 calibration cards per channel. For WRITEDB if more are needed, increase the size of arrays CAL1, CAL2, CAL3, CAL4. If more calibration cards are needed for WRITEDB2 only the arrays CAL1 and CAL2 need to be increased. The dB value of a data interpolation between the two values, i.e., if data value lies between the I^{th} and $(I-1)^{th}$

calibrations, is determined with the following equation:

$$\text{DB value} = \frac{(DB(I) - DB(I-1))\ (\text{DATA VALUE} - CAL\ (I-1))}{CAL(I) - CAL(I-1)} + DB\ (I-1)$$

The dB tapes represent data taken over a period of several hours. Some of the input signals were real data, some input represents noise, or periods when the recorder was not operating. By performing certain analysis on the entire file of data, those portions which represent valid data can be selected for more detailed study.

Two programs, SCINT36 and CH4ONLY were set up to perform analysis on a file of data tapes in increments of 1 1/2 minutes and give a running statistical summary of the results. Included in the output are means, median, S_4 values, autocorrelation values, cross correlation maximum, time of maximum cross correlation and velocities. SCINT36 is used to process 2 channels at a time. CH4ONLY does 1 channel at a time and hence no cross correlations are calculated for obvious reasons.

SCINT36 reads in data from a binary tape, two channels at a time. Certain parameters must be initialized, to allow for differences in some of the input tapes. ICHWD = no. of words per record, ICHRO = no. of data records between time records, LAST = number of words to be read in for a 1.5 minute sample, DIST = distance between the two channels. For tapes from Ancon, Peru, the distance between channels 1 + 3 = 366 m., between 1 + 2, DIST = 244 m., between channels 2 + 3, dist = 122 m., CH4 is not normally crossed with the other three. (For program CH4ONLY, DIST is not used.)

A 1.5 minute sample is read into an array. The program then calculates the median of the dB values, and the 98 percentile point. It then converts the dB values to power for the remainder of the analysis. The program finds the values for the mean, S_4, the deviation of S_4, the Nakagami-M value, determines the time for which the autocorrelation = .5 and 0, the standard deviation of the times, the maximum cross correlations and the time of the maximum cross correlation, the variance of the cross correlation, and the velocity of the cross correlation. The algorithm used to calculate the median mean of power, 98 percentile in dB, S_4, standard deviation of S_4, Nakagami-m autocorrelation, standard deviation of autocorrelation, cross correlation, standard deviation of cross correlation, and velocity are described in the final report, Mathematical Analysis and Implementing Software for Physical and Engineering Data.

These values are printed out with suitable headings, for each 1.5 minute segment of data, starting with the beginning of the tape and ending at the EOF, or at a preset segment number.

Program CH4ONLY is similar to above except that only one channel of data is read from the input tape and output does not include cross correlation information.

The data calculated above can be analyzed further by using either the program ANAL4 or PWRSPC. In ANAL 4, data is analyzed in 1.5 minute segments (either one or two channels at a time). Program has the capability of calculating and plotting any of fifteen combinations of the following: data input autocorrelation, cross correlation, power spectrum, probability density function, cumulative density function along with a Chi-squared goodness of fit test, message fade rate, and reliability. In addition, header information on the plots include the starting tape times, channel number, S_4 value, and satellite ID.

The program PWRSPC finds the power spectrum of a stationary time series, and plots it on a semi-log graphs. The program also has the option of calculating the message reliability and fade rate, and plotting them. The message reliability is plotted on a log-log graph, the fade rate is plotted on a semi-log graph. Examples of the three kinds of plots appeared in Section 8 of the final report of May 17, 1977 - May 16, 1978, Mathematical Analysis and Implementing Software For Physical and Engineering Data.

In ANAL4 , the data is read in from tapes. Two data cards are read in to select the analysis options. The program calculates the mean, and S_4 values, regardless of the options chosen, since these values are used extensively. The data is plotted after being read, with the axes TIME vs dB. The autocorrelations and cross correlations are calculated from power values, and plotted with a Y-axis from 0 to 1, and an X-axis representing elapsed time. Autocorrelation is plotted up to 16 second time lag, cross correlation up to 60 seconds lag for each channel, 0 time lag being in the middle.

The CDF and PDF functions are calculated from the data in dB, and plotted on a probability graph. If this option is selected, a Chi-squared goodness of fit test is done, on the power values. The results, along with the cumulative distribution, are printed out. The power spectrum is calculated using a rectangular window, and then smoothing the results by taking a ten-point running average of the results

The program PWRSPC finds the power spectrum of a time series of N points by performing a fast Fourier transform on the series, smoothing the output by taking a ten point running average and then plotting the logarithm of the power spectrum on a linear Y-axis, labeled from 0 to -100 dB, versus a log scale of the frequency.

Data is read in from one channel only, and up to fifteen minutes of data can be processed. The position, amount, number of plots and certain parameters are read in from data cards. The options to skip the sine curve check and the message reliability-fade rate subroutine, must be set at the beginning of the program. If the sine curve check is done, it will calculate the power spectrum of a sine curve which has been divided into 8192 parts. The power should peak at a frequency of 4096. The axes of the power spectrum are set for: X between 10^{-2} and 10^{2}, Y between - 100 and 0 dB. These axes are the same as in program ANAL4, for direct comparison purposes.

The fade rate is a plot of the cumulative distribution of the duration of the scintillations at certain levels, in this case at -2, -4, -6,

-8 and -10 dB levels, relative to the median. The message length chart is a plot of the message length versus the percentage of messages received perfectly at that dB level, and is related to the fade rate as follows:

$$R(d,L) = L - \frac{R}{N}, \text{ where } d = \text{dB level, } L = \text{length of message,}$$

N = number of possible messages in a time T, R = number of messages that are blocked.

If a measure of the stationarity of the input signal is desired, the program PRINTS4 can be used. This program compares the value of S_4 calculated over a 15 minute data sample with the values of S_4's calculated on 10 consecutive 1.5 minute sub-samples. Data is read until a 15 minute sample of data is stored in an array. Every sixth point is extracted for the calculations. First the value of S_4 for the entire sample is found, then S_4 for each 1.5 minute segment is calculated. The differences between the fifteen minute S_4 and each 1.5 minute S_4 are found and summed. The standard deviations of the 10 small segments and of the large one are found, along with the respective means, and third and fourth moments. The formulas used are as follows:

$$S_4 = \frac{\overline{X^2} - (\overline{X})^2}{(\overline{X})^2}$$

$\overline{X}$ = mean of series

$\overline{X^2}$ = mean of square of series

Conversion from dB to power is:

$$X_i = 10^{.1*(DB_i)}$$

The moments are:

$$M_2 = \frac{1}{N} \sum_{i=1}^{N} \left(X_i - \overline{X} \right)^2$$

$$M_3 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^3$$

$$M_4 = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^4$$

The standard deviation of the S_4's is:

$$STD(S_4)_j = \frac{S_4}{N}\left[\frac{(M_4 - M_2{}^2)}{4M_2{}^2} + \frac{M_2}{(\overline{X})^2} - \frac{M_3}{(M_2 * \overline{X})}\right]$$

These values are all stored in arrays, an. are printed at the conclusion of each segment. The program repeats the 15-minute increment until an EOF is reached.

Another program was written to test a selected sample of data for certain distributions. The program tests a sample (sampling is done for both a rate of 6 samples/sec., and a rate of 1 sample/1.5 seconds) for a certain distribution by setting up equiprobable cells, and then counting how many points fall into each cell. The sum of the differences between the expected value of the cell, and the actual value provides the basis for the Chi-squared goodness of fit test. The program matches the real distribution of data into cells with the theoretical distribution, using the IMSL subroutine GFIT. If the real data matches the assumed distribution, then the data will be equally distributed among the K cells, the difference between the actual count and the expected count in each cell gives the values for the K components of the Chi-squared statistic. The sum of the vector elements is the resultant Chi-squared statistic.

This program contains the subroutine necessary for comparing data with four different distributions: Nakagami-M, Gaussian, Rayleigh, and linear, but any number of different distributions can be tested. The one that provides the best fit is the one that is used most often. For the data in this project, the Gaussian distribution and the Nakagami-m distributions give the best fit, and are the ones used. There is also the option of producing a printer plot of the cumulative distribution.

From the observed signal levels, the amplitude and rate characteristics of intense scintillations were analyzed. It was observed from the distribution that the data closely resembles a Rayleigh distribution which is a special case of the Nakagami-m distribution for m = 1. The autocorrelation and power spectrum define the fading rate and give a basis for evaluation of different data. The cross correlation gives a means of evaluating space diversity at similar frequencies. Confidence bands are included to give credence to the reliability of the results.

7. IONOSPHERIC DATA ANALYSIS

INITIATOR: A. SNYDER PROJECT NO: 4643 PROBLEM NO: 4927

BACKGROUND

The objective of this effort was to determine the accuracy of the current recommended techniques for forecasting and specifying ionospheric parameters. Predicted f_oF2 values were compared with observed data. The world-standard method of specifying ionospheric parameters was used in conjunction with the analysis task undertaken. It was desired to improve upon the forecasting accuracy by incorporating corrections based on measured data. Data obtained from the DMSP satellite was the data used in this analysis.

Radio noise of ground-based origin can be observed by a satellite borne receiver orbiting above the F2 region. The observed noise is a function of received frequency, satellite position and local time at the sub-satellite point.

Defense Meteorological Satellite is equipped with a swept-frequency HF noise receiver. The receiver provides measurements of radio noise of terrestrial origin every 100 KHz in the frequency range of 1.2 to 13.9 Mhz.

The receiver continuously sweeps through the 128 frequency channels in 32 seconds. Thus, a value of the radio noise in a 100 KHz location is obtained once every 32 seconds. Because the satellite contains a recording system, data is obtained throughout the entire orbit.

Satellite is basically in a sun-synchronous morning/evening orbit. By analyzing data from successive morning or evening orbit, global maps of the noise intensity can be obtained at the satellite altitude. Work has been done in order to obtain world maps of noise intensity so as to assess the morphological behavior of the noise at satellite altitude as a function of frequency. Our studies have been concentrated on the frequencies of 1.5 to 13.5 Mhz in steps of 0.5 Mhz at both the dawn and dusk orbits. The following is a description of the functional steps in obtaining world noise maps and of the available software.

SOFTWARE AVAILABLE

The Air Force Global Weather Central processes several types of special sensor data including the data from the swept-frequency HF noise receiver. The DMSP tapes have been processed by GWC on the UNIVAC 1110 system, where the word length is 36 bits. In general, each DMSP tape contains dawn orbits of one day and dusk orbits of another day. These tapes are sent to AFGL on a regular basis. Under previous contract, these data tapes were read. Data was extracted and processed.

DATA EXTRACTION

Due to the changeover to the new system, the extracting software for DMSP data was not operational because of two of the system character and bit routines. Work was done to modify the software to extract the data from DMSP tapes without using the routines. This extracting program is necessary for the calculation of daily average count rates which is the first step in obtaining monthly DMSP count averages. It could be used to derive an improved f_oF2 prediction algorithm.

In the extracting program, RUSH, data is being extracted from the DMSP tape with the aid of subroutine BFF. Subroutine BFF reads 672 UNIVAC 1110 36-bit words into CDC 6600 60-bit words and unpacks each 36-bit words into 60-bit left-justified words. Program handles data blocks of 392 words at a time where BFF results in 1120 60-bit words in each call to it.

A check is included in the program to determine if this block is a data set (one ephemeris group plus 60 seconds of sensor data) or other information (i.e., information blocks or zero-filled words). Words 4 and 9 are checked to see if they contain the Julian days indicating the block is a data set. Program will search through the data block to see if the two words are present later in the data set and if so, a new data is defined. This program disregards any data where the geographic latitude is below -70.0° or above 70.0°, a new data set would be defined in this case. A flag system is presented in the program to prevent any duplication of data. There is a flag for each minute of day. If there is a duplication, a new data block will be defined, disregarding the duplication. The noise data is present in word 34 and every sixth word after. Each word contains the channel indicator, n, along with count rates at channels n-3 thru n.

This program calculates daily count rate averages for the frequencies of 1.5 to 13.5 Mhz in steps of 0.5 Mhz. There are either one or two 28x72x25 arrays generated from the program illustrating the averages as functions of latitude, longitude, and time depending upon whether one or two days of data are available. The program only handles two days worth of data at a time. Indices of the arrays are determined by geographic coordinates (5° x 5° block) and time (dawn or dusk).

Each data word is examined and the count rates for the frequencies 1.5 to 13.5 in steps of .5 Mhz are stored. To save space, the number of observations and cumulative sum are stored together in one word for dawn and dusk at each coordinate for the first day of interest. For the second day of interest, the three indices, count rate, and the local time are written on a second tape because of large core memory are ready accumulated with one 28x72x25 array. For the first day, where count rate was from dawn orbit, the values at destinated indices would be in the last 6 digits of the word and if another count rate is encountered for same indices, it will be packed the same way, ID1 (KA, KB, KF) = ID1 (KA, KB, KF) + 10000 + XZ where ID1 is array with KA, KB, KF being latitude, longitude and frequency indices. 10000 is added to indicate another observation has been encountered and XZ is actual count rate. For count rates from the dusk orbits, the equation would be:

ID1 (KA, KB, KF) = ID1 (KA, KB, KF) * 10E9 + XZ * 10E5

placing the number of observation and count rate in first 6 digits of the words. Count rates below 10 will be disregarded in the calculation.

The procedure continues until the end of file is reached. At this point, the average count rates are calculated for the first day. Each data word is unpacked and new data word is created for the same indices containing the number of observations and average multiplied by 10. Then the array is printed and stored on a permanent file.

For the second day, the second tape is read and a new 28x72x25 array is set up for both dawn and dusk orbits simila: to the set-up for the first day. The averages are calculated in a similar .anner as with the first day,also stored and printed. The flag systems for both days are also stored on a permanent file with the Julian days.

AVERAGING

The program BRUSH is the intermediate step in the extraction of SSIP data and generation of monthly and/or seasonal averages. This program combines multiple, up to five, data extractions provided by program BRUSH. Further, it may combine this summary with existing combined data. This program provides input for processing and analysis performed by program DIS, which calculates average count rates for frequencies of 1.5 to 13.5 Mhz in steps of 0.5 Mhz for both dawn and dusk as functions of latitude and longitude.

The program BRUSH reads a card determining which days and data sets to be averaged. Data, usually stored on permanent files, are then combined. Input for this program is the output generated by program RUSH. The output of BRUSH is the input for program DIS. The program combines input data, resulting in packed output values. That is, the number of points included and the average for both dawn and dusk values for a corresponding $5^{\circ} \times 5^{\circ}$ geographical block and frequency are determined and saved. Twenty-five frequencies are examined. The program usually uses permanent files as input, however, it could be used to access data on magnetic tapes. A file is designated to contain, where applicable, data from previously averaged data sets and is combined with the new averages. A utility file is set up to check for existing data on tape. That is, time is not spent averaging zeroes.

NOISE STATISTICS

Background statistics are calculated in the program DIS from the averages calculated in BRUSH. Further, this program produces digital contour listings as well as to allow for a floating criteria for background noise.

The primary function of this program is to produce concise monthly or seasonal noise statistics. The program uses as input the output generated by program BRUSH. These values contain both the number of points and the average count values observed in a $5^{o} \times 5^{o}$ geographical block per frequency. Program DIS unpacks this information per tape input and if applicable combines with other months or periods of time. The input values are unpacked by the equations ID = input values/1000000 = number of points.

KD = Input value - ID *1000000 = average values.

The number of data sets, either one or three, to be examined is determined via an input card parameter. The resultant averages are displayed in a 28 x 72 matrix per frequency, and time of day. Averages as a function of latitude and longitude in 5^{o} increments are also calculated. Means, standard deviations and number of occurences below-specified count levels are calculated and printed after the average count per frequency. Levels of interest vary per frequency. The total number of non-zero (zero meaning no data available) geographical block count averages is determined. The minimum level of interest is 30. The maximum level is the multiple of five above the criteria set to obtain background. That is, if the criteria is set at 80% and 1000 non-zero points are available, statistics are calculated for 30, 35, 40, etc. until statistics for 800 occurences have been calculated. Background is then calculated. Background noise is defined as the average plus one standard deviation of the values below the criteria value (always a percent) times the number of non-zero average values available.

A "poor man's contour" is then calculated and displayed. The background noise is subtracted from the average noise and displayed in a 28 x 72 matrix. The value 0 means the average is 4 counts above background or lower, 1 means the average is between 5 and 9 counts above the background, 2 means the average is between 10 and 14 counts above background, etc., a blank indicates no data available for the block.

ANALYSIS AND PROCESSING

During the past contract, F19628-78-C-0157, the monthly averages of count rates have been calculated for the months o. September through November for the frequencies starting at 1.5 continuing to 13.5 Mhz in steps of 0.5 Mhz for both dawn and dusk passes. All of the daily averages as well as monthly averages have been stored for future analysis. Background statistics were also calculated for the three months. Seasonal statistics were produced for the fall of 1977; i.e., September thru November.

Solar zenith angles and local hour angles were calculated for representative orbits of the data (from three months) examined. These zenith angles were compared to the worldwide zenith angles to see how close the data on the DMSP tapes were taken to sunrise and sunset. The satellite appears to have passed above the vicinity of ground-based ionosonde stations twice a day, once around dawn (0500 - 0700 local) and once around dusk (1600 - 1800 local). Zenith angle, can be derived from the following equations:

$$ZA = \text{arc cos} \{ \sin(\text{lat}) \sin(\text{decl}) + \cos(\text{lat}) \cos(\text{decl}) \cos(\text{SLHA}) \}$$

where: lat = geographic latitude
decl = solar declination
SLHA = solar local hour angle

The solar local hour angle, SLHA, was determined as follows:

$$SLHA = (15. \times UT - 180.0) \times \frac{\pi}{180^\circ} \quad \text{where UT = Universal Time}$$

Solar declination values are dependent upon the time of the year for which the calculation is done. Local hour angle was derived by:

$$LHA = \arccos \left[\frac{\cos (90.83) - \sin (LAT) - \sin (DECL)}{\cos (LAT) - \cos (DECL)} \right]$$

where LAT = geographic latitude
DECL = solar declination

Other tasks have been accomplished during the performance of this effort. The following are examples of the subtasks:

1. Print-plots and CALCOMP plots of count vs. frequency were generated for three time periods of day 241. One plot was generated for each set of 128 frequencies and approximately 350 of these plots have been plotted.

2. CALCOMP plots displaying the overall distribution of counts with respect to the 128 frequencies were plotted for the same three time periods as in 1.

3. Copies of J-prep files of September and November DMSP tapes were generated. Sample processing program along with description of it has been provided with these copied data.

8. ATMOSPHERIC OPTICAL PROPERTIES ANALYSIS

INITIATOR: E. SHETTLE PROJECT NO: 7621 PROBLEM NO.: 4946, 4948

BACKGROUND

AFGL/OPA is operating an atmospheric optics field station in northern Germany as one part of the NATO OPAQUE research program. The data collected by this station was merged with meteorological data and stored on a random-access mass storage device. A series of programs were written to process data available and generate statistics on the gathered data bases.

SOFTWARE AND ANALYSIS UPDATE

The statistical analysis has been performed on the following transmittance recorded by two transmissometers (Barnes and Eltro), horizontal illumination, temperature, dew point temperature, relative humidity, water absorption, wind speed, and wind direction. The analysis included plots of the data values, plots of the various averages, frequency distribution, histogram, cumulative frequency distribution, scatter plots, autocorrelations, and cross correlations. The statistics program has an option to perform this analysis on either 1 or 3 month(s) of data. This effort is composed of modifications and additions made to software package which is used to help in the analysis of atmosphere optical properties.

In addition to these changes, the entire set of plots were rerun because of errors in the original data submitted to ASEC. An algorithm was added to obtain plots of the wind speed and direction with respect to time.

The wind direction and magnitude are plotted with time as the abscissa. Wind direction is done on a ± 400 scale, because some measurements traversed 360^{o}. These values were repeated by subtracting 360^{o}, and making breaks in the curve to avoid giving an impression of a quick change in wind direction. Two sets of plots are generated, one for the original data, and one for the smoothed data for the temperature data.

A check on the reliability value of all data items was added to the existing program to prevent the plotting of the values with a reliability value of 0 or 1. Excluded from the plotting was the Eltro transmittance and Barnes transmittance after February 24th. Additions were made to improve the scaling of temperature, dew point temperatures, and relative humidity. Relative humidity is plotted as a dark line for ease of identification. The remaining changes to the plotting routine affected the algorithm for obtaining plots of the luxmeter data. The plotting capabilities were expanded including data for all hours instead of the hours between 0400 and 2000. This was done by expanding the arrays and modifying some of the indices. Algorithms were changed to check the maximum value of each hour of each day and compare it with the minimum value of each hour of each day and suppress plotting of these 2 data points if they are greater than one order of magnitude apart, and print the values at output along with their associated time. The 20, 50, and 80 percentile value of each hour was obtained and 3 lines were drawn on the plot which joins these percentile values for each hour. The luxmeter values for each hour of the data set are compared with the average for each hour, and values that are one order of magnitude greater than the average are printed out. Since the data set is saved in an array in its entirety, this stop is easily accomplished by first obtaining the average, and setting up a do loop to compare each data point.

The Barnes transmittance data was found to be in error, and this was corrected by multiplying the data in the original files by these correction factors and creating new mass storage files.

After the changes were incorporated, it was discovered that the scatter plots caused the drum to scour because of the high concentration of points in one area. It was decided to modify the software to generate CRT plots of the scatter plots. This involved changing the library subroutine calls. The following is a description of the software developed to assist in the Atmospheric Optical Properties Analysis.

SOFTWARE DEVELOPED

The statistics program first opens the mass storage files and initializes plot parameters. All other parameters are then initialized, and MP is set to 1 or 0 depending upon whether one month of data or three months of data will be analyzed. The first two input cards are then read and printed for visual identification of each case. The first card contains the number of the input arrays for the data-item. Two such data-items are processed from one card. The second card is used to determine what analysis will be performed.

The 21 cell values are then calculated, and the first record of the first month of data is brought into memory. The station ID, year, month, day and hour are extracted and printed. A do loop is set up to process those data values desired. First, its reliability value is checked and if it is 0 or 1, the value is not used in the analysis. If it is greater than 1, then the time in fractions of days is saved, the reliability value is converted to proper units and saved in an array. Frequency distributions are calculated using these values. This process is set up for two such input parameters, read from the input card. If only one value is desired, the second group is set to zeroes. Another record from the main storage file is read, and this loop continues until a month of data has been processed. Cumulative distributions for the month are then obtained and printed. The plot routines are then initialized. Depending upon the analysis chosen, the following are the items which can be performed upon the data set always on a monthly basis: (1) plot the maximum and minimum value for each 4 hour interval, (2) obtain plots of illumination, (3) plot the temperature and relative humidity, (4) plot the wind direction and wind speed. A check is then performed to see if monthly statistics or statistics for the entire 3 months are desired. If the latter are selected, more data is read from the next 2 files and the above procedure continues until all 3 files have been processed, and the entire 3 months of data saved in memory. Otherwise, the statistical analysis continues for just 1 month.

The following is a list of the options which can be selected for either 1 month or 3 months of statistics: (1) plot frequency distributions, (2) plot cumulative distribution, (3) plot histogram, (4) scatter plots, (5) plot autocorrelation, (6) plot cross correlation.

After the full 3 months have been processed, a check is made to see if the last data card has been read. If there are more input cards, all parameters are re-initialized, and a new case is read from input (2 cards).

Visible transmittance versus time plots are obtained from data taken with Eltro Transmissometer. The transmittance (T) is calculated from the extinction value (σ) as follows:

$$T = e^{-0.5\sigma}$$

This value is then multiplied by 100 to give percent for the ordinate scale; the day of month is used for the abscissa. Maximum and minimum values are obtained for each four-hour interval and two curves for each plot are generated showing the maximum and minimum value for each monthly period. Breaks in the plots indicate that data is either missing for that time, or bad. A histogram of percent of observation versus visual transmittance (0.5Km. path) can be obtained for the three month period or monthly. The ordinate scale is percent, the abscissa is visual transmittance in 5% intervals. The total number of observations is indicated above the plot. An array is set up which counted up the values which fall within each 5% incremental step. Along with the histogram above, a plot of the cumulative probability versus visual transmittance is obtained.

Transmittance versus time plots for TR spectral bands are obtained from data taken from Barnes Transmissometer for specific wavelength ranges of 3-5 μm, 8-12 μm, and open (4 μm). These plots are otherwise similar to the Eltro Transmissometer plots including the method of obtaining maxima and minima values. Histograms and cumulative probabilities are also obtained in a manner similar to that described above.

Scatter plots can then be obtained for either three months or one month with the abscissa and ordinate going from 0% to 100% transmittance for the following: 3-5 μm versus visible transmission, 8-12 μm versus total water vapor along viewing path, and 8-12 μm versus total water vapor along viewing path.

Water vapor for the scatter plots above are computed from the measured dew point temperature, μ, where,

$$\mu = 0.05\ p(t_D)$$

$$p(t_D) = \frac{216.68\ EV}{CV(T + 273.16)}$$

where EV = Goff-Gratch formulation for saturation vapor pressure over water,

CV corrects the compressibility of water vapor for deviations from the ideal gas law,

T = temperature in ^{o}C.

Day illumination versus time plots are obtained from horizontal luxmeter data. Ordinate scale of lux is a log scale from 10^{-3} to 10^{5} with the abscissa going from 0000 hours to 2400 hours for each month. There are, therefore, as many as 62 data points for a particular hour in a month. (Maximum and minimum value)

For each hour of the day, the 20%, 50% and 80% levels are found and a line is drawn connecting these values from hour to hour.

For night illumination, histograms of percent of observations versus lux are obtained. The abscissa is from 10^{-5} lux to 10 lux. Data values are taken from 1800 hours to 0600 hours for the time period. A corresponding cumulative probability versus lux may also be plotted.

The above algorithms, including plots, form one software package where the user specifies on an input card the data-item he wishes analyzed, and, on the next, the analysis he wishes, by means of a 1,0 code. The problem encountered in this package was development of a flexible enough plotting routine which would allow for the various scales and labels. As a result, two subroutines were created, one for linear plots, and one for semi-log plots. Labels and plot headers are created by obtaining the proper label from a data block by indexing with the data-item number in the data-base.

The next plot available is an overlay of three plots, temperature, dew point temperature and relative humidity. The latter is computed from the following:

$$RH = 100 \frac{e(t_d)}{e(t)} \frac{p-e(t)}{p-e(t_d)}$$

where t_d = dew point temperature

t = temperature

$e(t_d)$ = saturation vapor pressure at temperature t_d

$e(t)$ = saturation vapor pressure at temperature t

p = air pressure = 1013 mb

RH = relative humidity

To obtain a better representation of good data, the standard deviation is obtained for the temperature of each month. The difference between the data value and the average value about the point for a time of ± 2 hours is calculated and compared to three times the standard deviation. If the difference is greater than the latter value, the data value is discarded for plotting purposes.

Two ordinate scales are used for this plot, one for temperature $-20^{\circ}C$ to $20^{\circ}C$ and one for relative humidity from 0% to 100%, and the abscissa is the day of the month.

In addition to the above plot, the data is smoothed by taking a three point running average such that:

$$X_i = \frac{(x_{i-1} + 2x_i + X_{i+1})}{4}$$

where $i = 2, \ldots n - 1$, and

n = total number of points.

DATA COLLECTION

In addition to the statistical package, other software was generated in support of the total analysis. Software was written to extract atmospheric optics and meteorological data from the German weather stations. The met data on tape are unpacked and stored on new tapes in preparation for merging with the atmospheric optical data stored on mass storage files. The met data is merged with the atmospheric optical data using time values as indices. After this data has been merged, it is then ready for the statistical analysis.

Two programs are involved in this package. Program LOADIN, extracts the meteorological data from the tapes creating a new tape of the devised values; program OPFILE merges the meteorological data with the atmospheric optics data and creates new mass storage files.

Program LOADIN first reads the option card for print or punch, then it reads the beginning and ending dates and time for extracting the data. Subroutine METTS is then called to initialize parameters and to select the tower of interest. The input parameters are printed, and the first record is read. The records are decoded and the time is checked. If it is within the dates expected, it is saved until three records are obtained. These three records are printed and written onto tape until all the derived data has been extracted.

Program OPFILE opens the mass storage files and reads a tape record and a mass storage file record. The data and times are checked. The meteorological data is inserted into words 69 thru 77. Processing continues in this manner until an end-of-file is reached on the tape, or until all days of the month have been processed on the mass storage file.

Each month is processed separately.

Once the data is merged, updated, etc. copies of the data are supplied to the Central Data Bank. In connection with this activity software was written to prepare, in a specified format, data for use at the Central Data Bank.

This program uses as input a mass storage file and the file contains (85x24) words per record. The output tape consist of 990 characters/block with records not exceeding two blocks. The block is blank filled if a full record cannot fit. The unpacking program is used to check the output tape to make sure it is in the previously defined format.